CRWR Online Report 05-09

ArcGIS and HSPF Model Development

by

Nathan William Johnson, B.S.E.

David R. Maidment, Ph.D

Lynn E. Katz, Ph.D

August, 2005

CENTER FOR RESEARCH IN WATER RESOURCES

The University of Texas at Austin

J.J. Pickle Research Campus

Austin, TX 78712

This document is available on the World Wide Web at:

http://www.crwr.utexas.edu/online.shtml

Acknowledgments

To my parents… without your constant support and encouragement, I don't know where I would be.

I would like to thank my advisors, Dr. David Maidment and Dr. Lynn Katz for their encouragement and support. To PBS&J and CUAHSI, for supporting this research; to AQUA TERRA Consultants in Decatur, GA, for help in big and small things; to those at CRWR; for making the last two years fun and enlightening; and to God, for giving me a sharp mind and placing me in a position to use it; many thanks.

ArcGIS and HSPF Model Development

by

Nathan William Johnson, M.S.E.

The University of Texas at Austin, 2005

Supervisors: David R. Maidment, Lynn E. Katz

Abstract

The Hydrologic Simulation Program in Fortran, (HSPF), is a widely used computer model used to simulate hydrologic and water quality processes. Geographic Information System (GIS) tools from the BASINS system are widely used for creating new HSPF models, but operate within the ESRI ArcView environment. New ArcGIS software is an upgrade from their ArcView and ArcInfo predecessors and offers a robust data management framework and a common environment for the development of custom applications. This thesis presents a set of existing and custom tools that can be used for creating HSPF models in the ArcGIS environment. In addition, an ArcGIS system for preparing NEXRAD precipitation data for input to HSPF models provides an example of how ArcGIS tools and data structures can aid in HSPF modeling after initial model development.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AML – ArcMacro Language

API – Application Programming Interface

ATC – AQUA TERRA Consultants

BASINS – Better Assessment Science Integrating Point and Non-Point Sources

BRWMC – Bexar Regional Watershed Management Coalition

CUAHSI – Consortium of Universities for the Advancement of Hydrologic Science, Incorporated

CRWR – Center for Resources in Water Resources

DEM – Digital Elevation Model

.dll – Digital Link Libraries

.dss – Digital Storage System (File)

EPA – Environmental Protection Agency

ESRI – Formerly "Environmental Science Research Institute," now only ESRI

GenScn – GENeration and Analysis of Model Simulation SCeNarios

GIRAS – (Landuse dataset from the USGS)

GIS – Geographic Information Systems

GISWR – GIS in Water Resources (Consortium)

.grib – GRIdded Binary (File)

HIS – Hydrologic Information System

HRAP – Hydrologic Rainfall Analysis Project

HSPF – Hydrologic Simulation Program in Fortran

IDM – Interface Data Model

IMPLND – Impervious Land Segment [HSPF Operation Module]

LIDAR – LIght Detection And Ranging

LULC – Land Use / Land Cover

MRLC – Multi-Resolution Land Characteristics

NED – National Elevation Dataset

NEXRAD – NEXt Generation Weather RADar.

NHD – National Hydrography Dataset

NLCD – National Land Cover Dataset

NOAA – National Oceanic and Atmospheric Administration

NRCS – Natural Resource Conservation Service

NSF – National Science Foundation

NWS – National Weather Service

PERLND – Pervious Land Segment [HSPF Operation Module]

RCHRES – Channel Reach [HSPF Operation Module]

RWMS – Regional Watershed Management System

SAC-SMA – Sacramento Soil Moisture Accounting

SCS – Soil Conservation Service

SSURGO – Soil Survey Geographic Database

STATSGO – State Soil Geographic Database

SWAT – Soil and Water Assessment Tool

TACC – Texas Advanced Computing Center

TMDL – Total Maximum Daily Load

.uci – User Control Input

USGS – United States Geological Survey

WCS – Watershed Characterization System

.wdm – watershed data management

WMS – Watershed Modeling System

WRAP – Water Rights Analysis Package

Chapter 1 Introduction

1.1 BACKGROUND

Hydrologic modeling concepts have existed for hundreds of years; however, computer applications to hydrologic modeling became popular only after the late 1960's. Many of the early computer programs for hydrologic modeling have stood the test of time and become inimitable tools for hydrologists and engineers (Singh 2002). Even though computer technology and information science have progressed significantly over the past 40 years, the overwhelming familiarity and reputation of legacy models such as the Army Corps HEC-1 and HEC-2 hydrologic and hydraulic models, and the USGS's HSPF water quality model continue to make them indispensable tools for hydrologic studies.

HSPF (Hydrologic Simulation Program in Fortran) is a widely used water quality model based on a watershed hydrology model developed at Stanford University in the late 1960's. HSPF has capabilities to simulate both hydrologic and water quality processes on the land surface and in a river system (Bicknell et al. 2001). It has been used to simulate water quality in thousands of studies over the past 30 years and is one of the models recommended by the EPA (Environmental Protection Agency) for developing TMDL's (Total Maximum Daily Loads) (EPA 2005a).

Though recent advances in computer technology have not replaced HSPF and other legacy models, they have offered many benefits to the field of hydrologic modeling. The use of Geographic Information Systems (GIS) has permeated almost every field in the natural and social sciences, offering accurate, efficient, reproducible methods for collecting, managing, viewing, and analyzing spatial data.

Geographic Information Systems provide hydrologists with powerful tools to store, view, and analyze information about the environment through which water flows. GIS systems do not inherently have the hydrologic simulation capabilities that complex hydrologic models do, but they are capable of calculating many of the parameters that hydrologic models require (Whiteaker 2004). GIS data are widely used in the field of hydrologic modeling for determining the direction that water flows over the land surface from terrain data. GIS data are also used to determine land surface and river channel characteristics. Layers of spatial information can be combined in GIS to analyze the spatial distribution of vegetative cover and soil characteristics.

In addition to their analytical capabilities, new GIS systems provide a powerful data management framework. GIS data structures provide a consistent, intuitive platform for organizing and analyzing relationships amongst spatial objects and information associated with those objects. GIS data are widely used to describe the physical environment and researchers from many different disciplines use GIS for data storage, analysis, and transfer.

ESRI is the most widely recognized GIS software company in the world. Over the past 5 years, ESRI has reengineered their software to be more compatible with Microsoft and other Windows-based software products. Their new software products, ArcGIS 8 and 9, are referred to simply as ArcGIS. ArcGIS software is an upgrade from the previous ArcView 3.x and ArcInfo 7.x products which operated separately from one another. The new ArcGIS software is designed to facilitate the development of custom applications which are compatible with other Windows-based programs (Whiteaker 2004). In addition, ArcGIS uses a data management structure that is far more robust than the ArcView 3.x system. Geodatabases and data structures within them are used by ArcGIS software products to provide a well-structured, intuitive framework for storing spatial data, tabular data, and relationships amongst data.

Though GIS software and data structures have improved greatly in the past few years with the release of ArcGIS, most GIS preprocessing for HSPF is still done using ArcView software. ArcGIS tools and data structures have many advantages over the ArcView system but they are not widely used in the development of HSPF models because no standard methodology exists. The goal of this research is to develop a standard methodology for using ArcGIS data and tools in the development of HSPF models.

1.2 MOTIVATION

The most widely used GIS tool for developing HSPF models is the BASINS (Better Assessment Science Integrating Point and Non-Point Sources) system. BASINS is a software package with several GIS and non-GIS components all designed to aid in environmental assessment. GIS components of the BASINS system operate within the ArcView 3.x environment and contain tools for preparing input data for the HSPF model. In the GIS environment, ArcView terrain processing tools are used to delineate stream networks and drainage areas, and landuse data are used to estimate land surface characteristics for HSPF modeling. WinHSPF, a non-GIS component of the system provides tools to creating HSPF model files as well as a user interface for editing and running HSPF models after their initial development. (EPA 2001)

Though the BASINS system provides a well-structured, convenient way of using GIS data to develop HSPF models, many GIS users are moving away from the ArcView 3.x software in favor of the newer ArcGIS products. One of the most popular applications of the new ArcGIS software to hydrologic modeling is the Arc Hydro data model and its associated set of tools. The Arc Hydro tools operate within the ArcGIS software and are widely used to perform common tasks such as delineating drainage areas and defining stream networks based on terrain data. These basic tasks are general and not tailored to any specific hydrologic modeling application. Consequently, results of the Arc Hydro terrain processing tools can be used for any hydrologic modeling application.

In addition to the analytical capabilities of the Arc Hydro tools, the Arc Hydro Data Model provides a robust structure for organizing and managing GIS data related to water resources. Though neither ArcGIS nor Arc Hydro inherently have the capabilities to prepare input files for HSPF, they have the capability of performing all the spatial analysis required for HSPF model development and provide a data management framework that is unmatched by other GIS systems.

1.3 OBJECTIVES AND SCOPE

The objective of this research is to bridge the gap that currently exists between ArcGIS and the HSPF model. Specifically, an ArcGIS methodology for preprocessing data for input to the HSPF model has been developed. In addition, a methodology for organizing and preparing time series data in ArcGIS demonstrates the utility of using ArcGIS data and tools after initial model development.

The ArcGIS HSPF Preprocessing methodology, presented in Chapter 4, uses existing tools from the ArcGIS, Arc Hydro, and BASINS systems as well as custom ArcGIS tools developed in this research. The methodology starts with GIS data for stream networks and drainage areas developed with Arc Hydro or from another source. Utilizing existing Arc Hydro data avoids many routine terrain processing tasks that would otherwise have to be done by a preprocessing tool. This data is combined with land cover characteristics using standard ArcGIS tools to define the areas to be simulated by HSPF. Based on information from GIS data custom ArcGIS tools write text files defining the structure of a new HSPF model. Finally, WinHSPF, a non-GIS component of the BASINS system, is used to create new HSPF model input files. The custom ArcGIS tools developed as a part of this research are not intended to be "production-level" tools for using ArcGIS to preprocess data for HSPF, however, the prototype toolset may prove useful for many applications.

The ArcGIS Timeseries Preprocessing methodology, also presented in Chapter 4, uses the Arc Hydro time series structure, a well-established ArcGIS data structure, to organize and prepare input time series files for HSPF modeling. Custom ArcGIS tools use GIS data to automatically update HSPF model files to read input time series from the appropriate data sets. The ArcGIS Timeseries Preprocessing methodology is an example of how using Arc Hydro and ArcGIS for HSPF model development places HSPF models in the powerful data management and analysis framework provided by ArcGIS. As a result of the application of the ArcGIS HSPF and Timeseries Preprocessing methodologies, HSPF model files including input precipitation datasets can be developed efficiently using ArcGIS tools and data structures.

1.4 DOCUMENT OUTLINE

This thesis is divided into seven chapters. The first briefly provides some background placing the project in the context of Hydrologic Modeling and GIS applications to Hydrologic Modeling. The second chapter discusses GIS and Hydrologic modeling and presents some of the most widely used GIS tools for developing HSPF models, including the Arc Hydro and BASINS in some detail. The structure of the HSPF model is also presented in detail and Chapter 2 concludes with a brief discussion of time series structures and important parties that are interested in ArcGIS applications to HSPF modeling.

Chapter 3 discusses the use of GIS in HSPF model development and presents different methods of configuring HSPF models. The BASINS ArcView 3.x HSPF Preprocessing methodology is presented in detail followed by a discussion of time series development for HSPF modeling.

Chapter 4 presents the conceptual framework in which the ArcGIS HSPF and Timeseries Preprocessing methodologies are developed. Similarities to and differences from the BASINS HSPF Preprocessing methodology are highlighted and the concepts used in the Timeseries Preprocessing are presented. Chapter 5 presents the Application Procedure used to implement the methodology. The chapter begins with a brief presentation of the programming libraries used in tool development, and instructions for installing the tools. This is followed by a presentation of the data used by the ArcGIS

HSPF Preprocessing system, including a thorough description of a geodatabase structure used to support the analysis. Each new and existing ArcGIS tool used in HSPF model development is presented. A similar presentation of the data, geodatabase, and tools used by the ArcGIS Timeseries Preprocessing system concludes the chapter.

The results in Chapter 6 present five applications of the ArcGIS HSPF Preprocessing methodology to watersheds in the San Antonio, TX area. In addition, the use of ArcGIS data and tools after initial model development is demonstrated through the application of NEXRAD (NEXt Generation Weather RADar) precipitation estimates to three HSPF models. Conclusions are presented in Chapter 7 including future work that could improve on the current system.

Chapter 2 Literature and Model Review

2.1 Spatial Data and Geographic Information Systems

Interactions between objects in the world are often linked to their physical proximity, and having a physical representation of their location aids in understanding how they relate to one another. Spatial data that provide information about the precise location of objects in the environment have proven extremely useful for the study of environmental processes. Specific to the field of hydrology, an accurate spatial representation of characteristics such as the length of a river, the slope of the land contributing to that river, and the physical proximity of specific areas of land to the river are all extremely important to understanding the processes of water movement in the environment.

In the past, spatial information was stored almost exclusively on paper maps. Paper maps are helpful for observing spatial relationships between objects, but performing any sort of complex spatial analysis by hand is tedious. Hydrologic parameters such as slope, distance, and direction can be inferred or measured from a map, but cannot be extracted quickly or automatically.

A Geographic Information System (GIS) is, in essence, a digital system for storing, viewing, editing, and performing analyses on spatial data. A GIS typically takes the form of computer software that is capable of organizing, analyzing, and communicating with spatial data. It can be used to create paper maps, analyze the connectivity of transportation systems, or calculate parameters for environmental analyses. ESRI is the most widely recognized software developer of GIS computer programs. ESRI software supports applications in many areas including cartography, transportation, financial services, government, administrative, and natural resources.

“GIS data” are files that store information for GIS software to read and analyze. The two most common types of GIS data are vector data and raster data. Vector data usually take the form of a tabular file structure where each row in a table represents a single spatial object. Columns in the table contain attributes of the spatial object. f

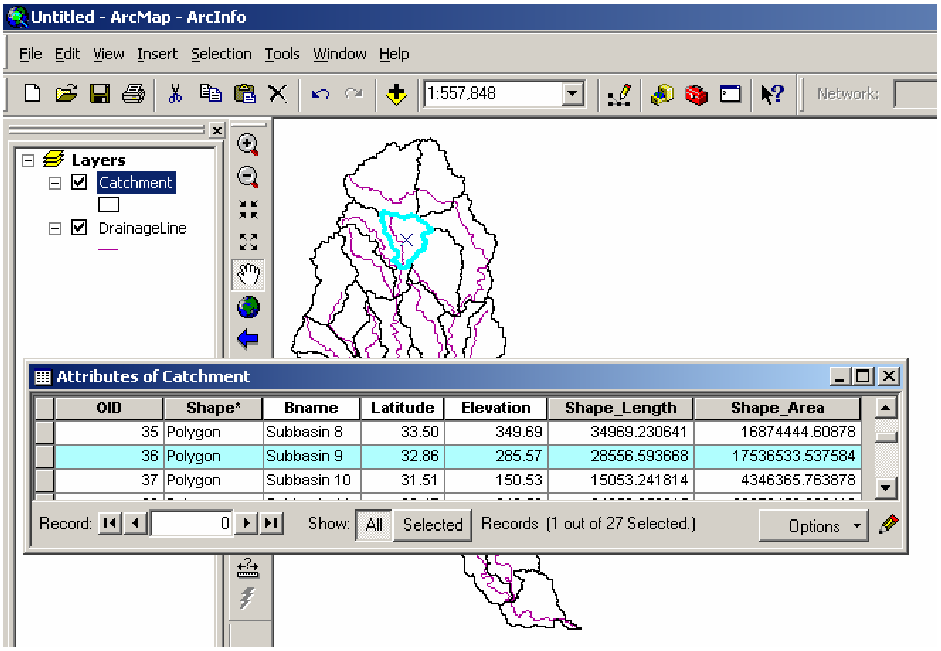


Figure 2.1 Vector GIS data shown in ESRI's ArcMap software.

The highlighted row in the "Catchment" GIS dataset has several attributes (columns in the table) including the latitude and elevation of each Catchment and a special attribute "Shape" which contains the spatial information shown in the map behind the table. The highlighted row corresponds to Subbasin 9, (also highlighted in the map) with an elevation of 285.57 meters and latitude of 32.86°N.

Raster data are typically used to represent a continuous spatial field of information. A single value representing the entity described by the raster dataset is stored at continuous, discrete intervals in space. Digital Elevation Models (DEMs) describe the elevation of the land surface and are often stored as raster datasets. Figure 2.2 is a screen capture from ESRI's ArcScene, a GIS software product for viewing data in three dimensions. It shows a raster dataset representing a DEM and a vector dataset of the catchments from Figure 2.1.

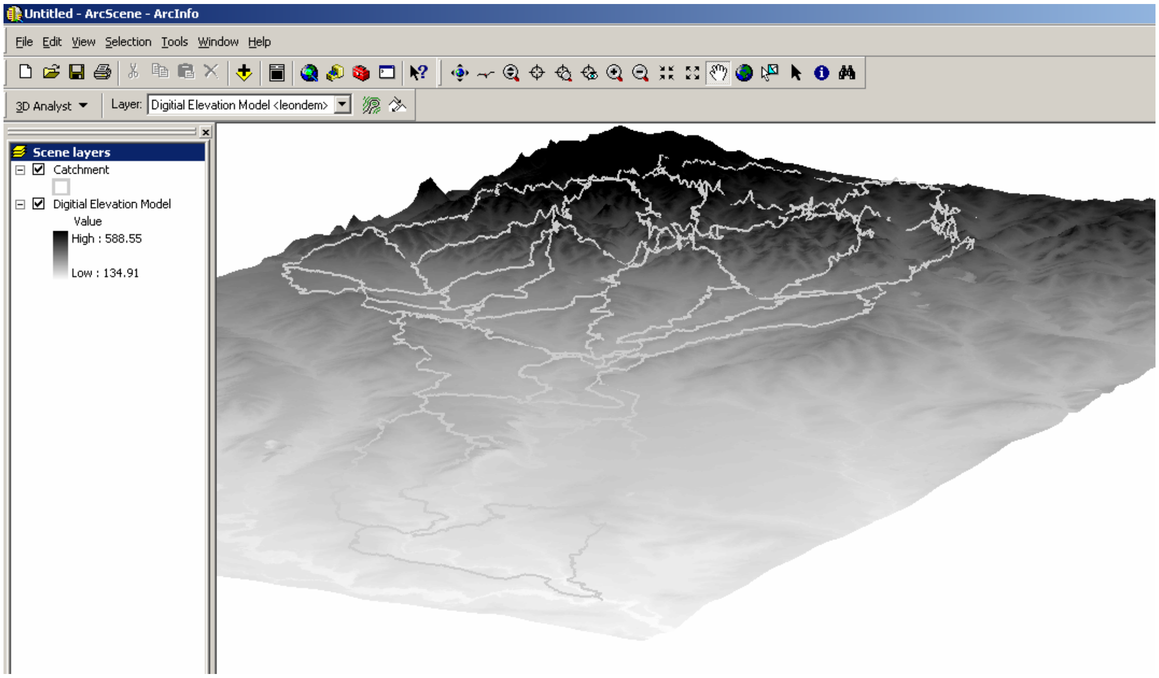


Figure 2.2 Raster GIS data shown in ESRI's ArcScene software.

GIS data are available for a wide range of applications and have been used in projects ranging from transportation system and urban planning to civil engineering and flood forecasting. GIS data for nearly all the streets and highways in the United States define connectivity and distance and support mapping applications, transportation planning, and Internet services such as "Mapquest.com." Spatial data about the locations of existing pipelines and utilities aid in planning and engineering the infrastructure to support city planning. Environmental data including river networks, land surface characteristics, DEMs, bridge locations, and locations of pollutant sources support analyses of surface water quality and quantity.

2.2 Hydrologic Modeling

Mathematical modeling is performed in almost every field that deals with processes occurring in time and space. Ecologists use statistical relationships to predict species distributions and populations; mechanical engineers use mathematical equations to predict the energy consumption of an engine during its design; and transportation engineers use complex algorithms to investigate the flow of traffic through an intersection. Hydrologic engineers use relationships amongst rainfall, infiltration, and runoff to estimate the flow in rivers for flood insurance and water availability applications.

The purpose of mathematical modeling is typically to characterize or predict conditions for which no observed data exists and to help understand the processes that are important for a system. Information about future population growth in an ecological community, the flow of traffic at a new intersection, or the future water quality in a river are all processes for which no observed data exist. Hydrologic models that investigate the influence of impervious area help hydrologists to understand the effects of urbanization on the response of a hydrologic system. Mathematical modeling of these processes allows interested parties to analyze the factors that affect the system response and make informed decisions in planning for future conditions.

Mathematical models can be classified in many different ways. One common approach is to classify them based on whether they are stochastic or deterministic. Almost all environmental processes are not completely understood and mathematical representations of these processes therefore contain a level of uncertainty. *Stochastic models* explicitly account for this uncertainty by using random fields to represent model parameters and variables. *Deterministic models* represent processes as being dependent upon precise values. They do not consider randomness in the processes they represent, and the same set of input values will always give the same set of output values. A deterministic model, HSPF, is the subject of this research, and the subsequent discussion will only cover deterministic models.

A further classification within deterministic models involves simplifications concerning spatial variability (Maidment 1988). A *lumped parameter* model does not explicitly account for spatial relationships between model parameters, inputs, or outputs. Lumped models usually have limited spatial resolution because they rely on spatial averages, but they reduce the complexity of a model significantly. *Distributed parameter* models account for spatial relationships amongst model variables and parameters. They typically use partial differential equations with respect to space to model the spatial distribution and influence of model variables. Later in this document, the terms 'lumped' and 'distributed' will be used to describe different configurations of HSPF models. Lumped and distributed configurations for HSPF models should not be confused with the traditional definition of lumped and distributed models. No matter what the configuration, HSPF is essentially a lumped parameter model. (Hydrocomp 2005)

Another simplification in mathematical modeling involves the time dependence of the processes represented. Many deterministic hydrologic models make the assumption of time-invariant processes, or *steady flow*. Steady flow means that the flow rate of water in the modeled system is assumed not to change over the duration of the model run. *Unsteady flow* models account for the variability of flow rate during the model run, which complicates calculations considerably. (Maidment 1988)

Another way that models are described is *empirical* vs. *physically based*. *Empirical models* use statistical relationships such as regressions between input to a system and the output response of a system without explicitly representing the physical processes that the system undergoes. *Physically based models* attempt to simulate the processes occurring in the system in as much detail as possible. For mathematical modeling applications in the field of hydrology, all the underlying hydrologic processes are not explicitly simulated because they are not completely understood, and even physically based hydrology models contain some empirical components.

An additional distinction in hydrologic modeling is that of *Continuous* vs. *Event Based*. An *Event Based hydrologic model* attempts to simulate the response of the landscape to a single rainfall event. It requires that initial hydrologic conditions of the landscape be known, and requires forcing (or input) data only for the duration of the rainfall event. *Continuous hydrologic models* simulate the response of water in the landscape over much longer times. Continuous models attempt to keep track of the hydrologic conditions in the landscape that affect rainfall-runoff response between storm events such as soil moisture. Initial conditions are also required for continuous models, however, model results become less dependent upon these initial conditions during longer simulations. (HydroComp 2005)

HSPF, the subject of this research, is a deterministic, lumped-parameter, physically based, continuous model for simulating the water quality and quantity processes that occur on watersheds and in a river network. It should be noted that all the mathematical model distinctions discussed above fall within the category of *abstract models*, as opposed to *physical models*. *Physical models* are typically downscaled models of the actual physical system. They reproduce processes on a reduced scale to observe the system response while *abstract models* represent a system with concepts, relationships, and mathematical equations.

In reality, all the processes that HSPF and other environmental models attempt to simulate are overwhelmingly complex, occurring continuously in three-dimensional space and time. If these processes were understood completely, a deterministic, continuous, distributed, physically based computer model could be developed to accurately predict the response of every point of a hydrologic system to forcing data such as rainfall, evaporation, pollutant deposition and transport. This model would require parameters that fully characterized the hydrologic and water quality processes at every point in the system.

Unfortunately (or maybe fortunately), all processes are not well understood, and HSPF and every other hydrologic and water quality model rely on different levels of spatial and temporal averaging to predict the response of a system. Mathematical models are developed to simulate processes as accurately as possible within the context of limitations presented by incomplete understanding of underlying processes and data availability. (Maidment 1988)

2.3 EXISTING TOOLS FOR GIS AND HYDROLOGIC MODELING

As early as 1993, the use of Geographic Information Systems (GIS) was growing in the area of water resources engineering. (DeVantier and Feldman 1993) Whether lumped or distributed, continuous or event based, deterministic or stochastic, all hydrologic models are based on an accurate representation of the landscape. Over the past 15 years, GIS data have proven to be a convenient, accurate, and efficient way to represent the environment through which water flows. In 1991, Maidment suggested four possible applications of GIS to hydrologic applications: hydrologic assessment, hydrologic parameter determination, hydrologic model preparation, and hydrologic modeling within GIS (Maidment 1991). Over the past 15 years, almost all applications of GIS to hydrology have focused on the first three of these applications. (Maidment 2002)

GIS systems have been used to map and visualize flood inundation areas and characterize areas with water quality problems (hydrologic assessment), (EPA 2001, Robayo 2005), to determine input parameters for many hydrologic models (parameter determination), (Bhaskar 1992, Olivera and Maidment 1999) and to prepare data for hydrologic models (hydrologic model preparation) (USACE 2005a, USACE 2005b, EPA 2001). This research deals primarily with hydrologic model preparation and parameter determination, and the most common tasks in using GIS for these purposes are outlined in the following paragraphs.

Hydrologic modeling usually is motivated by the need to predict the response of a river network to rainfall. Whether a lumped, distributed, physically based, or empirical model is used, the areas of land that flow to river segments (often referred to as "Drainage Areas" or "Catchments") are critical information. The most common use of GIS for hydrologic modeling is to automate the process of defining the area of land that flows to a river segment. DEM raster data represent the elevation of the land surface at consistent, discrete intervals. Because hydrologic processes are driven by gravity, water follows the path with the steepest slope over the land surface, and continuous elevation information can be used to determine the direction of flow at every point in the landscape. Once the direction of flow is determined, a contributing area threshold can be defined to predict where river channels will begin, or alternatively, contributing areas can be calculated for predefined river segments with a flow direction grid.

Another piece of critical information required for almost all physically based hydrologic models is the distribution of land surface characteristics within the drainage area of a river segment. Vegetation and soil type play an important role in defining the land surface characteristics, and GIS data describing vegetative cover and soil type are widely available. A lumped parameter model may use the average slope and dominant vegetation type of the drainage area in its calculations, while a distributed model may require information at many points within the watershed. In either case, GIS provides powerful tools to quickly and accurately calculate physical parameters required for model simulations.

The American Society of Civil Engineers published a series of papers in 2001 that provides a summary of current technology related to GIS and water resources applications. (Garbrecht et al. 2001, Ogden et al. 2001) This research concerns itself chiefly with parameter determination and model preparation, and a review of some of the current tools for these applications follows. Many GIS tools are available to aid in the development of specific hydrologic models. The Arc Hydro and BASINS tools, however, contain tools that are more generally applicable and are most relevant to this research. Arc Hydro and BASINS will be covered in some detail, and cited references for the other tools should be consulted for further explanations.

2.3.1 Arc Hydro Data Model

Arc Hydro is a customization of ESRI's ArcGIS software for water resources applications. Arc Hydro includes a data model and associated set of tools to support hydrologic modeling, but does not contain any simulation capabilities itself. The Arc Hydro data model includes a geospatial and temporal framework for storing information related to water resources. Arc Hydro tools are available to aid in populating the data model with attributes of hydrologic features. The Arc Hydro data model is intended to be only a framework for storing some of the general information necessary for hydrologic modeling in a way that can subsequently be customized to fit the needs of a specific model.

Tools included with Arc Hydro provide the functionality to perform many tasks that are common to all hydrologic model preparation. Delineating drainage areas and streams, calculating lengths, areas and slopes, and defining connectivity between system components are all tasks that are supported with the Arc Hydro tools. Figures 2.3 and 2.4 show the structure of the Arc Hydro data model and an example of the types of GIS data that are created by the tools and stored in the data model.

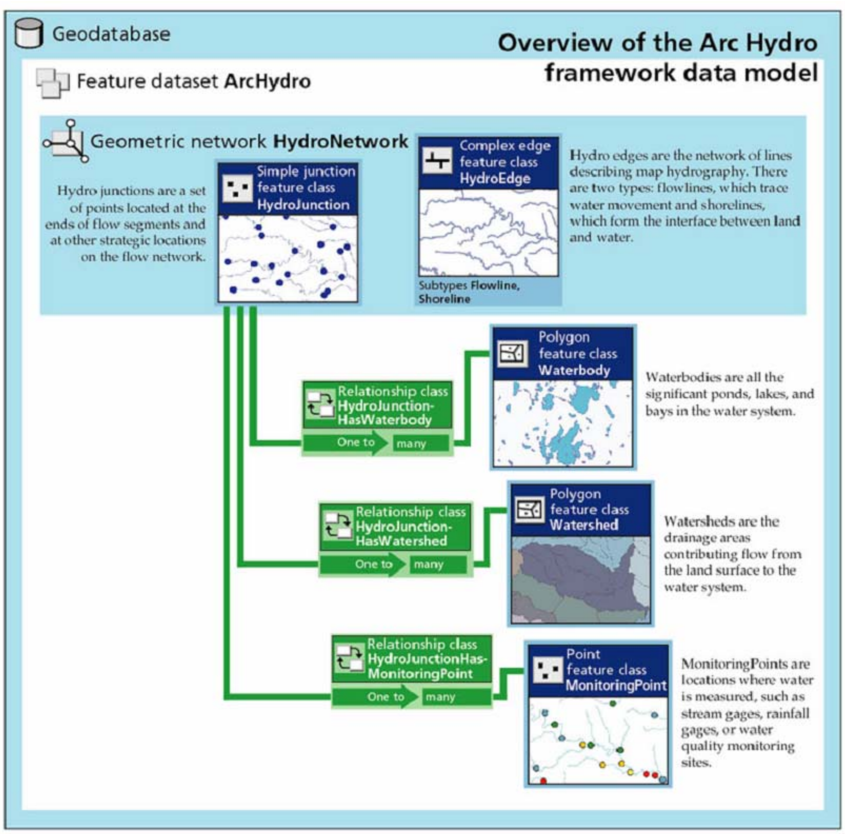


Figure 2.3 ArcHydro framework data model (Maidment 2002).

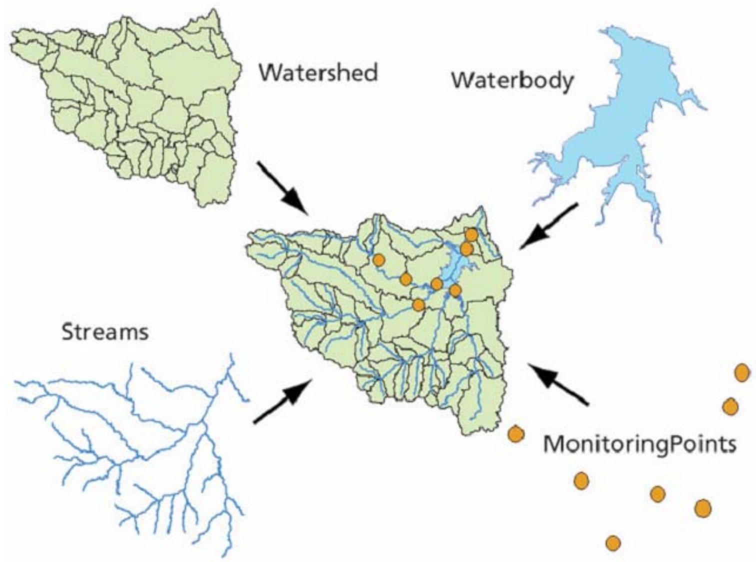


Figure 2.4 Example of data stored in the ArcHydro data model. Database structure defines connectivity between hydrologic elements. (Maidment 2002)

Figure 2.5 illustrates some of the tools available in the Arc Hydro toolset. Terrain Processing tools are available to delineate drainage areas using terrain data; Watershed Processing tools calculate areas, centroids, and longest flow paths for watersheds; and Attribute Tools manage data in the data model by finding downstream drainage lines or drainage areas. Network tools aid in preparing data to use the network tools available in ESRI's geodatabase structure.

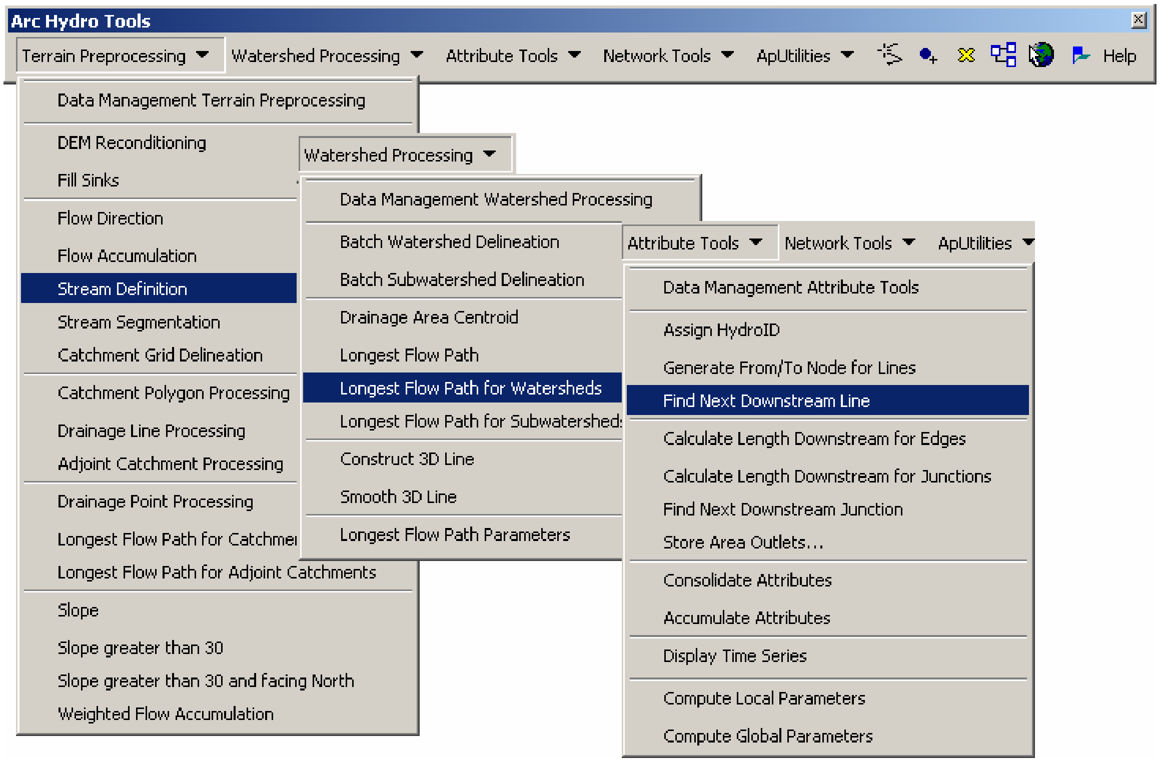


Figure 2.5 The ArcHydro Toolbar: Terrain Processing → Watershed Processing → Attribute Tools.

The research supporting the development of Arc Hydro was undertaken at the Center for Research in Water Resources (CRWR) at the University of Texas at Austin. The tools run within ESRI's ArcGIS proprietary software and the data model is built upon ESRI's Geodatabase data structure. Both the tools and data model are available free of charge from the GIS in Water Resources (GISWR) Consortium (http://www.crwr.utexas.edu/giswr/).

Arc Hydro and Hydrologic Modeling

Arc Hydro is helpful to the process of linking GIS data with hydrologic models for two reasons. First, Arc Hydro contains a toolset that can be used to calculate hydrologic parameters that are used by many different models. The Arc Hydro tools leverage the power of ArcGIS to perform spatial analysis with an interface that was designed specifically for applications to hydrology.

Arc Hydro contains tools designed to:

1) Calculate the area that drains to a river or gauging station based on digital elevation data.
2) Calculate downstream distances from points in a stream network.
3) Calculate parameters for watersheds such as centroid locations, longest flow paths, and slopes.

The second way Arc Hydro helps in preparing data for use with hydrologic models is by providing a framework for organizing and managing data. The Arc Hydro data model stores data in a general format that was designed to be independent of individual models, types of models, or applications. The data framework does not make assumptions about the scale of the data, where it came from, or what it will be used for. An attempt was made to capture essential relationships amongst hydrologic elements in the environment with intuitive representations that can be useful for a wide variety of applications. Tools to aid in organizing and managing data were designed to operate on this somewhat abstract hydrologic GIS data framework.

Specifically, Arc Hydro data model and data management tools are designed to:

1) Create and maintain unique identification information for each feature in a database.
2) Use these unique ID's to manage relationships between watersheds, land segments, river segments, drainage points, hydraulic structures, river morphology, and monitoring stations.
3) Provide a structure for storing and utilizing information relevant to networks and connectivity

Arc Hydro and IDM's

Though the Arc Hydro data model and tools do not inherently contain the capabilities to create, format, and prepare data for input to hydrologic models, the basic Arc Hydro framework has been extended and customized to fit the needs of several models. There are a few common parameters that are used by almost all hydrologic simulation models, such as area, slope, and distance, but each individual model typically relies on many additional parameters that are specific to the model algorithms. It is impractical to attempt to fit all of these model parameters into the relatively simple Arc Hydro database structure. The phrase "Interface Data Model" (IDM) is a term coined at CRWR to describe a database structure and associated set of tools that facilitate the transfer of information from model files to an Arc Hydro database (Obenour 2004). An IDM provides the structure to store model parameters in a way that is compatible with both the hydrologic model and the Arc Hydro data structure.

IDM's have been developed for both HEC-HMS, and HEC-RAS as a part of a project to regionalize the management of water resources information near San Antonio. (Obenour 2004, Roboyo 2004) The utility of Arc Hydro's abstract framework for storing water resources-related GIS data is apparent in this project. Eventually, an Arc Hydro geodatabase will be used to store the most accurate, up-to-date representation of the hydrologic environment. This standard GIS framework will include general representations of the physical landscape including high resolution DEMs, land cover and impervious area, river channels, and hydraulic structures. Models for water quality, stormwater, and flood protection will all be built upon and linked to this data and, as a result, will accurately reflect the current state of the environment.

All hydrologic modeling efforts are built upon some representation of the physical landscape through which water moves. However, model requirements and assumptions vary widely, and maintaining a consistent representation of the physical environment across modeling efforts is not simple. Maintaining a consistent set of GIS data in the Arc Hydro framework allows each model linked to the GIS data to have the same representation of the physical environment, such as river channel geometry, watershed lengths, distances, and areas, and landcover distributions.

Figure 2.6 shows how GIS data has been traditionally used for hydrologic modeling. Typically, a different set of GIS data is developed specifically for each individual modeling application. This data is used to create necessary model files and visualize the results of the model. Figure 2.7 shows how the Arc Hydro framework can be used to maintain a consistent representation of physical elements in the environment across modeling applications.

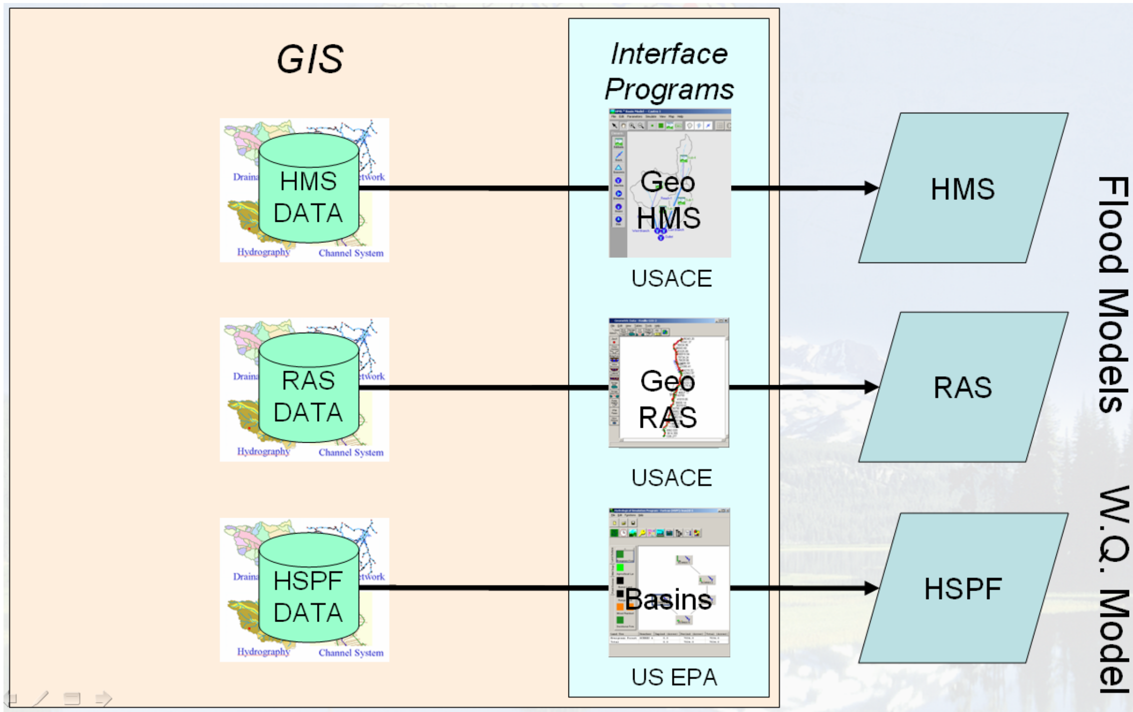


Figure 2.6 Traditional use of GIS for water resources applications.

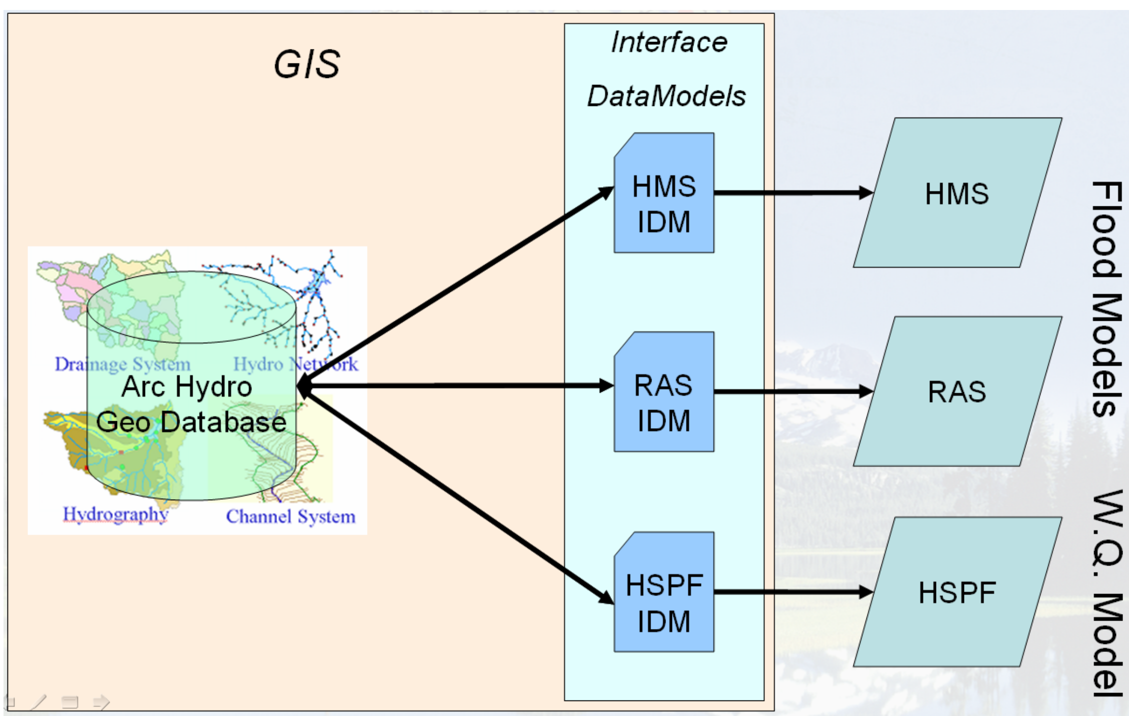


Figure 2.7 Arc Hydro: Framework for GIS and water resources applications.

2.3.2 BASINS and Model Preprocessing

BASINS

Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) was first released in 1996. It is described as a “multipurpose environmental analysis system for watershed and water quality studies” (EPA 2001) and has been supported by the Environmental Protection Agency’s (EPA) Office of Water for nearly 10 years. It is distributed free of charge from the EPA’s BASINS website (EPA 2005b), and supported by AQUA TERRA Consultants (AQUA TERRA 2005) through funding from the EPA. Though the software itself is available free of charge, some components run within ArcView 3.x, a proprietary software package from ESRI. An effort is being made to have next release of BASINS (v.4.0) operate as independently of proprietary GIS software as possible.

The BASINS system contains tools for hydrologic assessment, parameter determination, and model preparation. One component of the BASINS system, which is embedded in ESRI’s ArcView 3.x interface, provides access to a host of environmental data. Environmental data are distributed from the BASINS website for the entire country, and tools included in the BASINS system use this data to characterize the state of existing water quality on a regional scale.

In addition to hydrologic and water quality assessment tools, the BASINS system contains tools for hydrologic model preparation and parameter determination. BASINS tools are available for the most common tasks in hydrologic model preparation such as delineating watersheds and defining river segments. The GIS component of the BASINS system supports the preparation of data for use in several separate hydrologic, watershed, and pollutant loading models. Model parameters can be calculated for an instream water quality model, two watershed models, and a simplified Non-Point source pollutant model (EPA 2005b).

Though only the Non-Point source model has been integrated entirely into the GIS environment (the other models must be run outside of the GIS), BASINS provides tools to prepare input files for the models. In each case, data is prepared in a specific format within the GIS, and a tool is provided to create model input files using data stored in the GIS. The model files produced are used to run, calibrate, and perform analyses outside of the GIS environment. (EPA 2001) The process of using of BASINS and WinHSPF for creating of new HSPF models will be discussed in detail in Section 3.5.

WinHSPF / WDMUtil / GenScn

The relatively simple model input files created with GIS tools in BASINS are often not detailed enough to perform extensive analysis, and additional work is usually necessary. To aid in this process, separate programs are available to interact with the HSPF model input files, input time series data, and output time series results. The GenScn/WinHSPF/WDMUtil programs are components of the BASINS system that run outside of the GIS application. Though they are capable of reading shapefiles, their functionality does not require complex spatial analysis, and many of the files used by the programs are simple text files. WinHSPF provides a Windows-based interface to the complex structure of an HSPF model input file. All model parameters and even the structure of an HSPF model can be changed using WinHSPF. WdmUtil is a program for creating, managing, and performing analysis on .wdm (Watershed Data Management) files. Wdm files are the most common time series format used for HSPF modeling. GenScn (GENeration and analysis of model simulation SCeNarios) is designed to assist with the process of analyzing and managing the voluminous time series outputs associated with water quality model simulations (USGS 1998). It provides some simple mapping functionality and a framework for creating different scenarios of an HSPF model.

2.3.3 Other GIS and Hydrologic Modeling Applications

WRAPHydro

Wrap Hydro is a set of tools and a database structure designed to operate within ESRI's ArcGIS software and calculate the geospatial parameters necessary for the WRAP (Water Rights Analysis Package) model. The WRAP model simulates the management of water resources by assigning time and spatially varying priority information to discharge and withdrawal points in a river system. (Wurbs 2003) The WRAP Hydro data model divides its preprocessing procedure into three steps. The first step is gathering the data and loading it into an ArcGIS database format. The second step is preprocessing using Arc Hydro tools. The final step involves calculating the input parameters for the WRAP model including the connectivity of the system, downstream distance, average upstream area, average curve number, and average annual precipitation for each control point in the model. (Gopalan 2003)

HEC-PrePro, PrePro2002, GeoHMS

HEC-PrePro was originally developed at CRWR using Arc/Info's ArcMacro Language (AML) in 1997 (Hellwiger and Maidment 1997). The program automates the transfer of data from a GIS to the format necessary for input to the HMS hydrologic model. It contains a methodology for representing common hydrologic objects in a GIS and transferring that representation to elements for hydrologic modeling. CRWR-PrePro is a further development of HEC-PrePro that includes the tools for stream and watershed delineation and hydrologic parameter calculation (Olivera and Maidment 1999).

Geo-HMS, developed by the Hydrologic Engineering Center, includes cimilar capabilities to CRWR-PrePro, and it is the preprocessor used by the Corps of Engineers. PrePro2002 has been developed by Texas A&M University and Dodson and Associates, Inc. (Olivera 2005). It is similar to the previous preprocessors, but runs in the ArcGIS platform. It performs many of the general data preparation tasks such as delineating subbasins and streams, defining outlets, calculating areas and lengths, as well as preparing the data for input to the HEC-HMS model. All of these preprocessors run within proprietary ESRI software, but the tools are distributed freely from Dr. Francisco Olivera's website (Olivera 2005).

WCS

The Watershed Characterization System (WCS) has been developed with the support of EPA Region 4. It runs within the ArcView 3.x platform and was primarily designed to assist in the characterization of existing conditions in a watershed. The core program provides some analytic capabilities such as delineating watersheds and analyzing the physical and hydrologic properties of a watershed including soils, landuse, slope and water quality.

Additional functionality is added to the WCS by several extensions to assist in the modeling of sediment, mercury, stormwater, and non-point sources. GIS interfaces have been developed for the Storm Water Management Model (SWMM) and the Non-Point Source Model (NPSM). The Sediment and Mercury tools run within the GIS, and produce reports that have been used for TMDL development (EPA 2005c).

WMS

Though many of the tools discussed above run within proprietary software, the tools themselves are public domain, and freely available from various agencies and universities. The Watershed Modeling System (WMS) is one of the most widely used proprietary software packages for preparing hydrologic and hydraulic modeling data

Like Arc Hydro, the WMS does not contain any simulation technology in itself, but is described as a modeling environment for all phases of hydrologic and hydraulic modeling. It does not require the use of any external proprietary GIS software, but has been designed to work in conjunction with ESRI products to make use of their tools if desired. The WMS contains tools for general hydrologic data preparation as well as interfaces to a suite of hydrologic and hydraulic models. Supported hydrologic models include HEC-HMS, TR-20, TR55, HSPF, Rational Method, NFF, MODRAT, and GSSHA. GIS links to hydraulic models are supported for HEC-RAS, and CE-QUAL W2 (EMRL 2005)

In addition to providing access to individual models, WMS is capable of linking HEC-HMS and HEC-RAS to perform integrated hydrologic and hydraulic modeling. A stochastic approach to floodplain delineation is supported for examining the uncertainty in hydrologic and hydraulic modeling with HEC-HMS and HEC-RAS. The WMS software design is modular, and individual components are added for modeling can be added to the core functionality of WMS (EMRL 2005).

2.4 HSPF Model: Structure and Time Series

HSPF is a set of computer codes designed to simulate water quantity and quality processes occurring on the land surface and in stream systems. The algorithms implemented in the model are primarily based on a watershed hydrology model developed at Stanford in the 1970's. Since its initial release in 1980, HSPF has undergone many changes, most of which have added additional simulation capabilities. Though technology and science have progressed significantly over the last 20 years, the core algorithms have withstood the test of time and HSPF is still today the state-of-the-art in comprehensive watershed hydrology and water quality modeling software (Bicknell 2001).

2.4.1 Model Structure – Representation of the "Real World"

In designing a model to simulate processes occurring in the environment, a consistent view of the environment must be developed to provide the structure for mathematical modeling. The conceptual model used to develop HSPF consists of a set of constituents interacting with each other as they move through a fixed environment. HSPF constituents include water, sediment, and other chemicals that move through the environment. The fixed environment is, in reality, a continuous system, but for the purposes of mathematical modeling must be broken up into discrete parts.

In HSPF, the environment that constituents move through is described by two fundamental objects. “Nodes” are fixed points in space at which the value of a constituent can be measured. “Zones” are representative of a finite portion of the real world in which constituents reside and move through. These two abstract concepts form the building blocks upon which the HSPF representation of the “real world” is built. Nodes represent the endpoints of river segments, the location where two river segments converge, or any other specific location related to constituent movement. Zones are discrete pieces of the environment such as a length of a river, or a portion of the land surface or subsurface. Zones are associated with storage of a quantity of a constituent, and represent the smallest unit into which the environment is partitioned. In HSPF, collections of Nodes and Zones are called “Elements.” HSPF Model Elements are used to define the portions of the “real world” through which constituents move.

Physically, Model Elements represent the finite portions of the real world that, when linked together, represent the hydrologic environment through which water moves. Mathematically, Model Elements are a concept providing the structure upon which computer code simulates the movement and interactions of water and constituents. Figure 2.8 illustrates the concepts of nodes, zones, and elements.

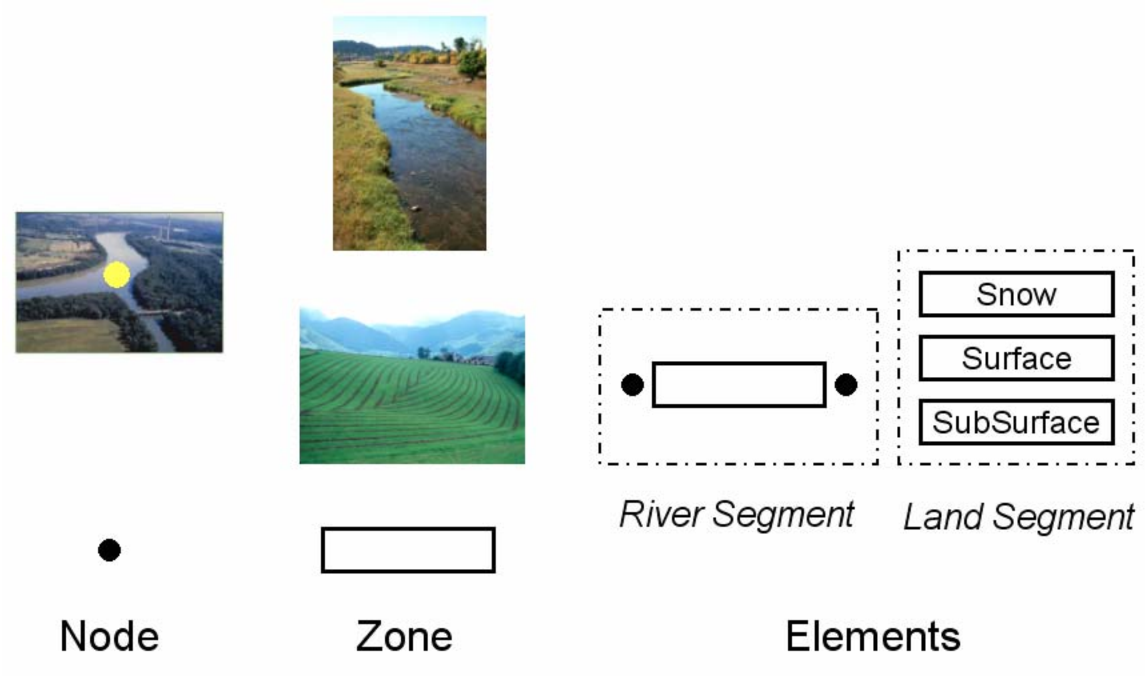


Figure 2.8 Elements of the "Real World" in HSPF.

The choice of grouping Zones and Nodes into Elements is made so that as much of the environment as possible can be represented by a single Element structure. There are two basic Elements Types in the HSPF model: River Segments and Land Segments. Regardless of its length, geometry, or slope, the same or similar processes occur in all rivers and HSPF uses Elements called River Segments to simulate the movement of constituents through the river network. Likewise, the same basic processes occur everywhere on the land surface, i.e. infiltration, runoff, groundwater flow; and HSPF uses Elements called Land Segments to simulate the movement of constituents over and through the land surface (Bicknell 2001).

River Segment Elements consist of a single Zone with two Nodes on each end. Land Segment Elements have one Zone to represent the surface and additional zones to simulate snowpack and the subsurface layers when necessary. In designing the structure of the HSPF model, it was decided that there were enough differences between

Impervious (no subsurface Zones) and Pervious Land Segments (1-3 subsurface Zones) to warrant creating different Element Types for each.

The three Element Types used to model all watershed and river processes in the HSPF model are Channel Reaches, Pervious Land Segments, and Impervious Land Segments. The same basic processes occur within each of these Element Types, and consequently, their mathematical descriptions share a common set of input parameters and use the same calculations to simulate the movement of constituents. Differences in the processes described by individual Elements within an Element Type, such as runoff from forested land vs. runoff from agricultural land, can be characterized by assigning different parameters to each individual Element.

Though these three types of Model Elements provide a starting point for dividing the environment into finite portions to be simulated by HSPF, the model is flexible, and a specific combination of Land Segment and River Segment Model Elements must be defined by the user in order to simulate a system of interest. This decision of how to 'Configure' HSPF Model Elements to simulate a hydrologic system is the subject of Chapter 3.

2.4.2 Software Structure

There are two major tasks performed by the HSPF model during a simulation. The first task is peripheral to the actual simulation and involves reading the input parameters provided by the user. The HSPF model obtains all the parameters required for simulation from a text file called the User Control Input (.uci). The input data must be read into the program memory so it can be used by the simulation part of the program. This task is carried out by a section of computer code called the Run Interpreter. The Run Interpreter and other sections of code called Utility Modules perform tasks that are auxiliary to the actual simulation. They read input parameters from the .uci file, read and write time series to external files, and manage the order of the simulation.

The second task of the model is to perform the actual simulation. The design of the code to simulate hydrologic and water quality processes is somewhat object-oriented and related to the structure used to represent the environment presented above. Because each Element Type is simulated by a common set of equations and parameters, the computer code contains subroutines, or ”Operation Modules,” specific to each Element Type. These Operation Modules contain the equations used to simulate the movement and transformation of constituents within a single Element.

Three separate Operation Modules are responsible for performing the calculations that simulate hydrologic and water quality processes, one for each Element Type. Because they are simulated by Operation Modules, Elements within an HSPF model (River Reaches, Pervious Land Segments, and Impervious Land Segments) are often referred to simply as “Operations” (Bicknell 2001).

The computer code within each Operation Module is further partitioned so that processes and groups of processes are simulated by different Sections of code. Individual Sections of code can be called so that if a user wants to simulate only water *quantity* processes in a river, the code for water *quality* need not be run. The Operation Module used to simulate Channel Reaches (referred to as RCHRES in the HSPF model) contains separate Sections of code to simulate hydraulic behavior, pH, temperature, advective transport, and a host of other water quality related processes. Similarly, Operation Modules for Pervious Land Segments (PERLND) and Impervious Land Segments (IMPLND) are partitioned into Sections to simulate overland flow, temperature, and the transport of many constituents. Figure 2.9 presents the structure of the HSPF software and how it relates to the “Real World” processes it simulates.

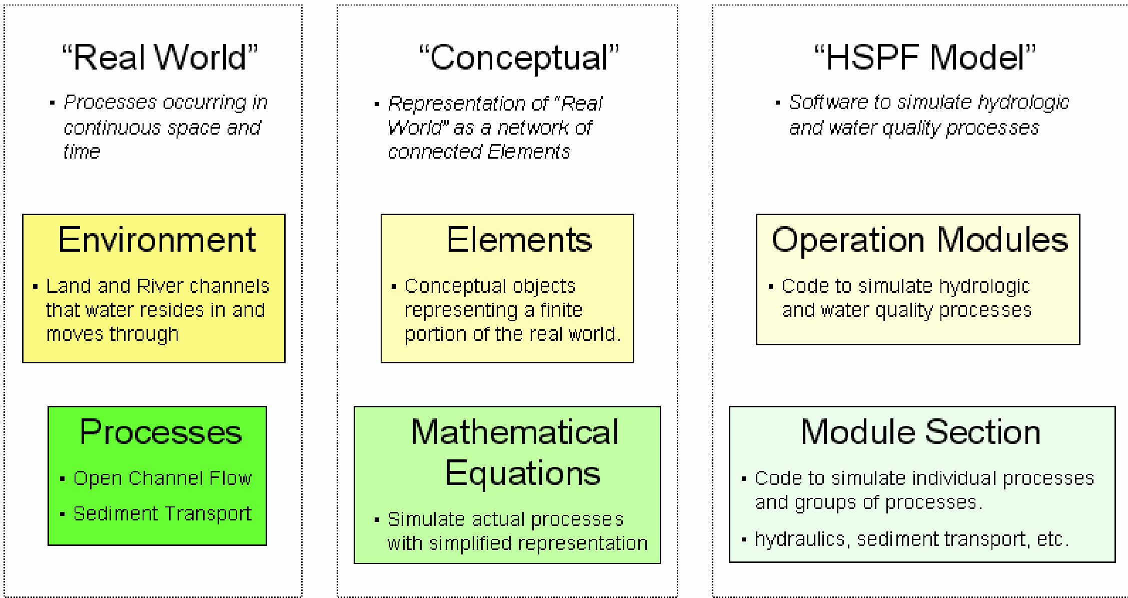


Figure 2.9 HSPF Representation of the "Real World."

The terms "Element," "Segment," and "Operation" all have a specific meaning when talking about HSPF and are used repeatedly in the remainder of this document. All three of these terms are used to describe the finite portions of the world which an HSPF model simulates. "Element" and "Segment" are used to describe the actual portions of the environment which HSPF simulates, while "Operation" typically refers to the model representations of these real-world objects.

2.4.3 Input Parameters and Time Series Management

The parameters that the HSPF model uses in simulations come from the User Control Input (.uci) file, a fixed format text file with a structure having some similarities to the code structure presented above. The .uci file is organized into groups of text lines called Blocks. There are Blocks of text lines for each Operation Module, PERLND, IMPLND, and RCHRES. Within each of these text Blocks there are Sections of text lines that correspond to the different Operation Module Sections. Each Section of text lines store the parameters used to simulate individual processes or groups of processes as mentioned previously. As an example, there is one Section of text lines within the RCHRES Block that stores the parameters used to simulate the hydraulic behavior of Channel Reaches. Likewise, there is one Section of text lines within the PERLND Block that stores the parameters used to simulate overland flow on Pervious Land Segments. Figure 2.10 illustrates the .uci file structure.

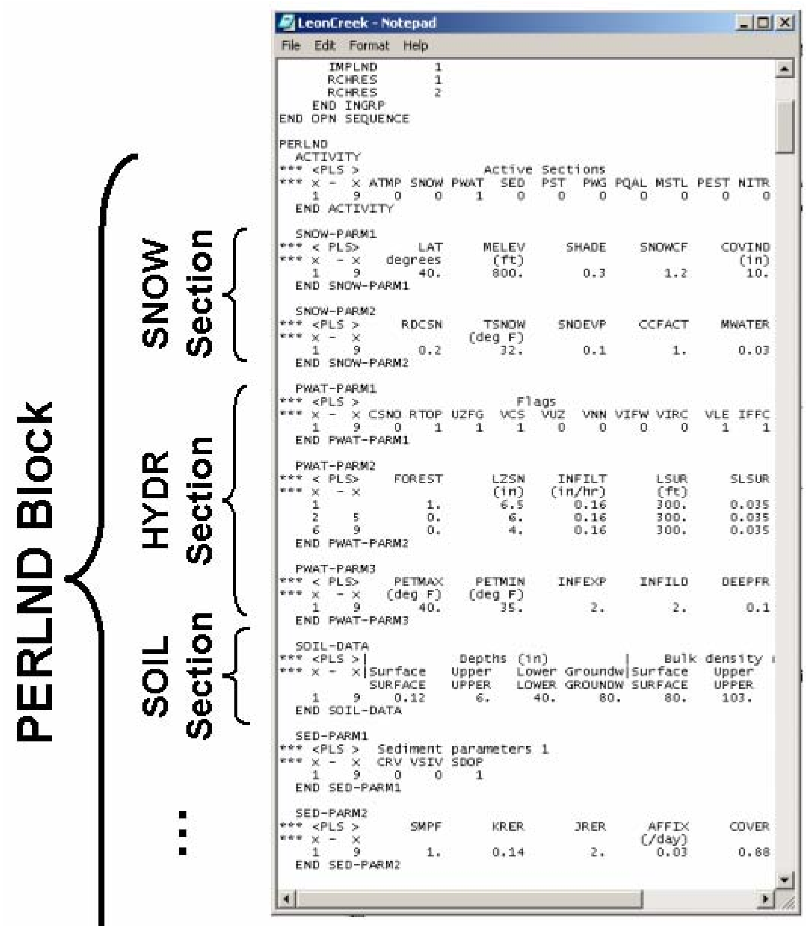


Figure 2.10 .uci file structure.

In addition to Blocks for each Operation Module, there are Blocks of text lines for some of the Utility Modules as well as Blocks to hold information that defines connectivity between Operations. The parameters in these Blocks do not participate directly in the simulation of hydraulic and water quality processes, but aid in the management of the simulation and point to hard disk locations of input and output time series.

As a continuous simulation model, HSPF must handle an immense amount of time series data. Complex HSPF models may have hundreds of Operations (PERLNDS, IMPLNDS, and RCHRES) and simulate processes on them for 5 years or more on an hourly time step. One hundred Operations times 10 processes times 5 years times 365 days times 24 hours is nearly fifty million values that must be passed amongst the proper subroutines in a complex HSPF model.

In addition to the sheer volume of data, simulations must proceed in the correct order when simulating an entire watershed and stream network system. The order in which the HSPF simulation proceeds is defined in a Block of text lines in the .uci file called the Operation Sequence Block (OPN SEQUENCE). In HSPF, a single Operation (simulating a finite portion of the real world) takes one or more input time series of arbitrary length and deterministically calculates the hydrologic and water quality response of the Element. Processes are simulated for the entire length of the simulation and the resulting time series is used (potentially with other time series) as input to the next downstream Operation. HSPF is designed to simulate the processes that do not depend upon any upstream conditions first. For instance: first simulations will be run for all the processes occurring on Land Segments that contribute to the most upstream River Segment. The output from these simulations (a time series of overland flow, nutrient load, etc.) will be used as input to simulate processes in the most upstream River Segment (hydraulic behavior, bacteria decay, etc.). The output from the most upstream river segment is subsequently used as input to the next downstream river segment.

The connectivity between model segments is defined in two Blocks of text lines called “Schematic” and “Mass Link.” Simulations on the land surface are computed essentially as a vertical water balance. Each Land Segment simulated by HSPF computes fluxes on an areal basis and contributes to the appropriate River Segment as defined in the Schematic Block. The Schematic Block defines the connections between Land and River Segment Operations and also specifies how much of the land area simulated by a Land Segment Operation contributes to each River Segment. For a complex HSPF simulation, many timeseries must be passed between Model Operations (i.e. water, sediment load, temperature, Dissolved Oxygen) and the Mass Link Block specifies explicitly which time series are to be passed over the connections defined in the Schematic Block.

While accomplishing this monumental task of organizing and managing the data for such a complex simulation process, the developers of HSPF designed the time series handling Modules in HSPF to be flexible. It is important to be able to retrieve information about the state of the system at any point during the simulation, however, writing all of the information to an external file would slow the simulation significantly and use massive amounts of disk storage space. The time series management Modules in HSPF avoid writing all the time series from a simulation to an external file by using internal memory to transfer time series that have a common length and timestep. This "internal swapping" of time series avoids the problems associated with outputting every time series, and is the default unless information from the user requests that time series be written to an external file. The HSPF model is capable of reading time series in several different formats; however, Watershed Data Management (.wdm) files are by far the most common input and output file format.

2.4.4 Concept of an Operation and HSPF Data Types

The HSPF model is built around the concept of an Operation. An HSPF Operation is a set of computer codes that simulate the hydrologic and water quality processes occurring in a finite portion of the environment for a certain length of time. At each timestep, an HSPF Operation performs four basic tasks to simulate the system response:

1) Read in *Forcing Data*,
2) Perform calculations using *Operation Parameters,*
3) Update the *State Variables* of the operation, and
4) Calculate *Output Data* for the operation.

Forcing data, such as precipitation or inflow to a river, is provided from an external file or an upstream Operation. Operation Parameters, such as infiltration coefficients or river lengths, are used by the Operation in conjunction with State Variables to predict the response of the modeled segment to the given Forcing Data. State Variables, such as soil moisture or River Segment volume, are updated at the end of a timestep and used as initial conditions for the subsequent time interval. Output Data, such as soil moisture or overland flow volume, are calculated by the Operation and sent to a downstream Segment or an external file.

The HSPF model structure is general enough that Model Operations can be linked together in many different ways to simulate a system. A single Land Segment may be used to simulate the entire area of a watershed, or the watershed may be broken up into hundreds of Land Segments to capture the spatial variability of hydrologic processes. Different methods for linking HSPF Operations together to simulate a hydrologic system are the topic of Chapter 3.

2.5 TIME SERIES STRUCTURE

Though there are literally hundreds of different specific timeseries formats used to store data for different modeling applications, this section will only discuss three general types of data structures most relevant to this research.

2.5.1 Grid Data Structure

Many spatially continuous environmental data sets are stored in a gridded file structure. NetCDF (Unidata 2005) and .grib (GRIdded Binary) (WMO 2005) formats have been developed by the atmospheric community to store observed or modeled data sets that represent a continuous spatial field. A truly spatially continuous data set would have values defined at every infinitesimally small finite portion in space. In reality, data for spatially continuous variables are typically stored in a grid format, where the values at the centers or nodes of the grid represent some sort of spatially averaged value. Usually these data sets, while considered continuous in space, represent only a single point in time. Temporally, the value could represent any type data; an average or cumulative value over the previous or subsequent timestep, or an instantaneous measurement at that exact instant.

A collection of these gridded data sets spaced at regular time intervals can be used to provide a continuous spatial and temporal description of conditions in the environment. The NetCDF file structure can store a collection of temporally consecutive spatially continuous data sets in a single file, while individual .grib files are necessary for each timestep. Whatever the details of the overall file structure, gridded data sets are used most often to represent spatially continuous data at a single point in time.

One benefit of a gridded data structure is that values are stored at regular spatial intervals, so that it is not necessary to store spatial information with every data value. Metadata at the beginning of gridded data sets provide all the necessary spatial information to locate a value in the data set for any point in space. For instance, Figure 2.11 shows the header information and the beginning of the data values from a .grib file (here extracted to ascii format) used by the West Gulf River Forecast Center (NOAA 2005a) to store NEXRAD (NEXt Generation Weather RADar) rainfall data (NOAA 2005b).

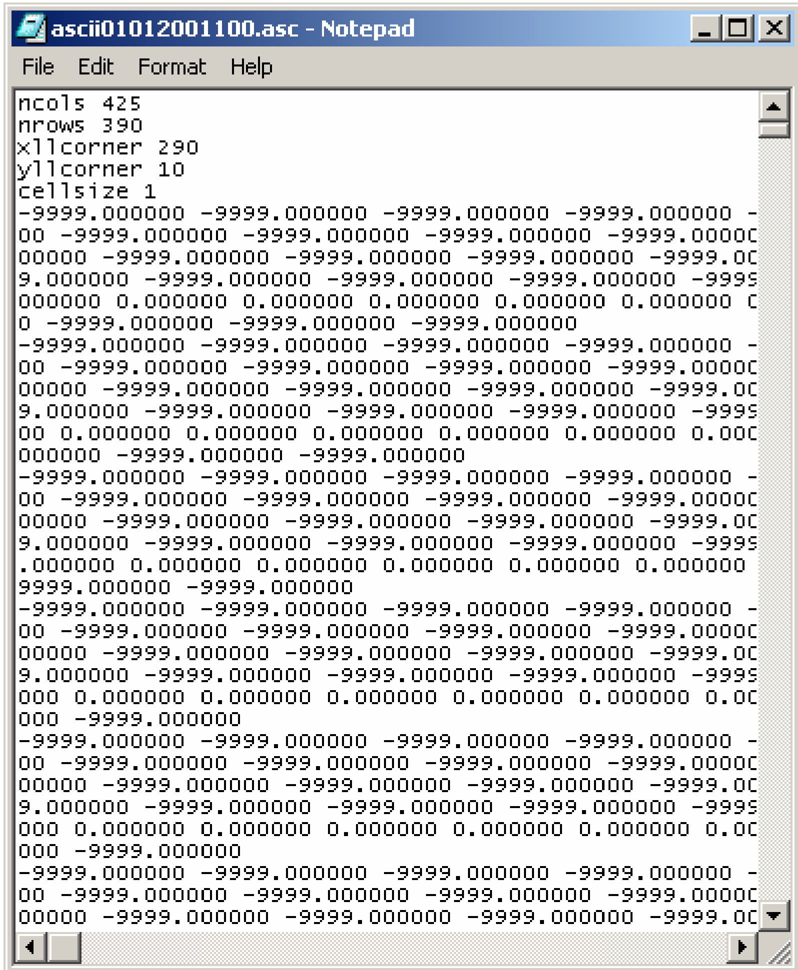


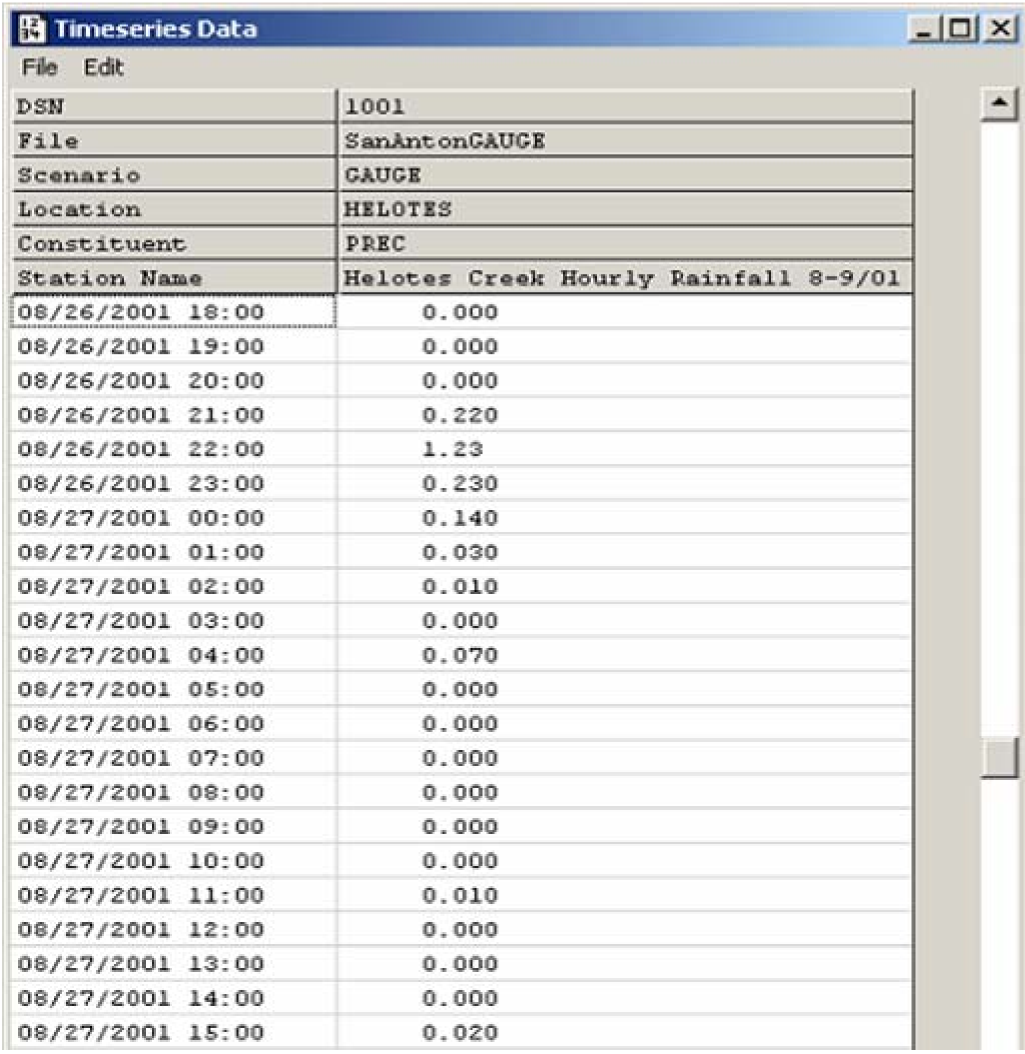
Figure 2.11 ‘Gridded’ file structure.

The data set values describe a spatial average (over the grid cell) of the cumulative amount of rainfall that fell on August 1, 2001 between 12:00 AM and 1:00 AM at each of about 164,000 HRAP (Hydrologic Rainfall Analysis Project) grid cells covering the southeastern US. Metadata at the beginning of the file give the spatial coordinates (in HRAP X and Y grid cell numbers) of the lower left corner of the data set, and the number of rows and columns contained in the data set. After the metadata, a simple list of about 164,000 numbers, delimited by a single space, provide the values for each of the grid cells described by the dataset.

This format is efficient in terms of disk space because it avoids storing spatial data with each value. The file structure, including the metadata at the beginning of the file and the subsequent list of values, inherently stores spatial information for each value in the dataset. However, in this type of gridded data structure, a value for each grid cell is required, regardless of whether a value was actually measured or is necessary at that location.

2.5.2 Modeling Time Series Structure

While grid data structures are used widely to store spatially continuous data, such as NEXRAD data and the results of atmospheric models, most hydrologic models require input data sets to be structured as a series of values at regular intervals describing individual locations. The .wdm (Watershed Data Management) (USGS 1995) and .dss (Digital Storage System) (USACE 2005) formats are examples of this type of timeseries data structure where metadata including spatial location, units, and type of data are followed by a list of values or date-time – value pairs. Each data set in this structure represents a timeseries of values at a single spatial location. Figure 2.12 shows an example of a .wdm file (here extracted to ASCII format) consisting of a list of metadata followed by time – value pairs for rainfall at a gauge near San Antonio.



DSN	1001
File	SanAntonGAUGE
Scenario	GAUGE
Location	HELOTES
Constituent	PREC
Station Name	Helotes Creek Hourly Rainfall 8-9/01
08/26/2001 18:00	0.000
08/26/2001 19:00	0.000
08/26/2001 20:00	0.000
08/26/2001 21:00	0.220
08/26/2001 22:00	1.23
08/26/2001 23:00	0.230
08/27/2001 00:00	0.140
08/27/2001 01:00	0.030
08/27/2001 02:00	0.010
08/27/2001 03:00	0.000
08/27/2001 04:00	0.070
08/27/2001 05:00	0.000
08/27/2001 06:00	0.000
08/27/2001 07:00	0.000
08/27/2001 08:00	0.000
08/27/2001 09:00	0.000
08/27/2001 10:00	0.000
08/27/2001 11:00	0.010
08/27/2001 12:00	0.000
08/27/2001 13:00	0.000
08/27/2001 14:00	0.000
08/27/2001 15:00	0.020

Figure 2.12 Modeling time series data structure.

This structure is distinct from the grid data structure discussed earlier, where a data set represents a single point in time at all spatial locations. Figure 2.13 illustrates the differences between the two types of timeseries structures using the space – time – variable domains. In both the gridded and timeseries structures, the data sets represent a series of values that all have the same variable type, but are continuous in either space or time.

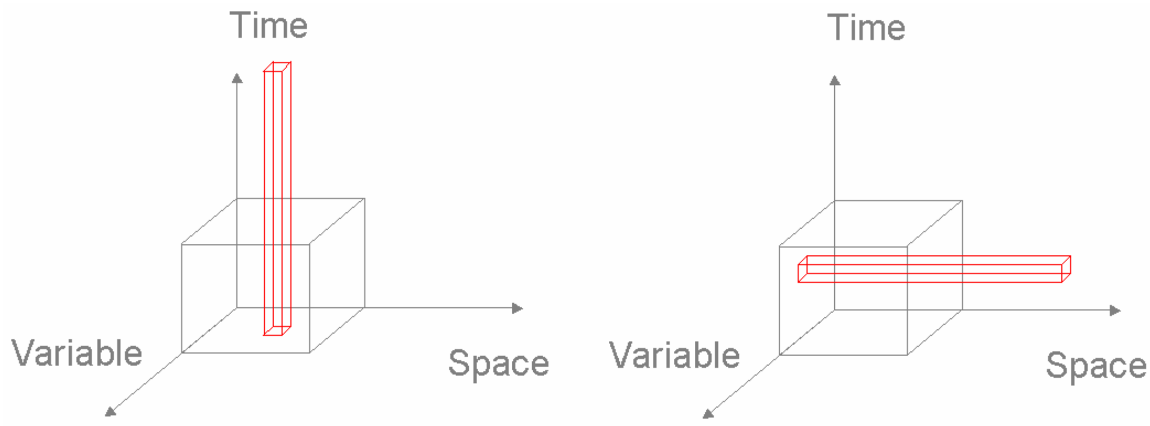


Figure 2.13 Grid and Time Series data structures in space – time – variable domains.

2.5.3 Arc Hydro Time Series Structure

The Arc Hydro data model is housed within the structure of an ESRI's geodatabase, and consequently has a structure different from either of those presented above. In the Arc Hydro format, a single table contains all the timeseries values in the database. Spatial information is present on each timeseries record in the form of a 'FeatureID' corresponding to a spatial feature in the geodatabase. In addition to having spatial data on each timeseries record, the Arc Hydro format also includes metadata on each record in the form of a 'TSTypeID' corresponding to a record in the TSType table. Figures 2.14 and 2.15 (Maidment 2002) demonstrate how timeseries are linked to geospatial features in a geodatabase.

Table
TimeSeries

TimeSeries is a single large table storing time varying attributes of the features.

Field name	Data type	Allow nulls	Default value	Domain	Precision	Scale	Length	
OBJECTID	OID							
FeatureID	Integer	Yes			0			HydroID of the feature described by the time series
TSTypeID	Integer	Yes			0			Identifier for the type of time series
TSDateTime	Date	Yes			0	0	8	Date and time of the time series value
TSValue	Double	Yes			0	0		Time series value

Table
TSType

TsType is an index of the types of time series data stored in the time series table.

Field name	Data type	Allow nulls	Default value	Domain	Precision	Scale	Length	
OBJECTID	OID							
TSTypeID	Integer	Yes			0			Identifier for the type of time series
Variable	String	Yes					255	The variable described by the time series, like streamflow
Units	String	Yes					255	Units of measurement
IsRegular	Integer	Yes		AHBoolean	0			Whether data regularly or irregularly measured by time
TSInterval	Integer	Yes		TSIntervalType	0			Time interval represented by each measurement
DataType	Integer	Yes		TSDataType	0			Type of time series data e.g. instantaneous, cumulative
Origin	Integer	Yes		TSOrigins	0			Origin of the time series dat

Figure 2.14 Arc Hydro Time Series tables.

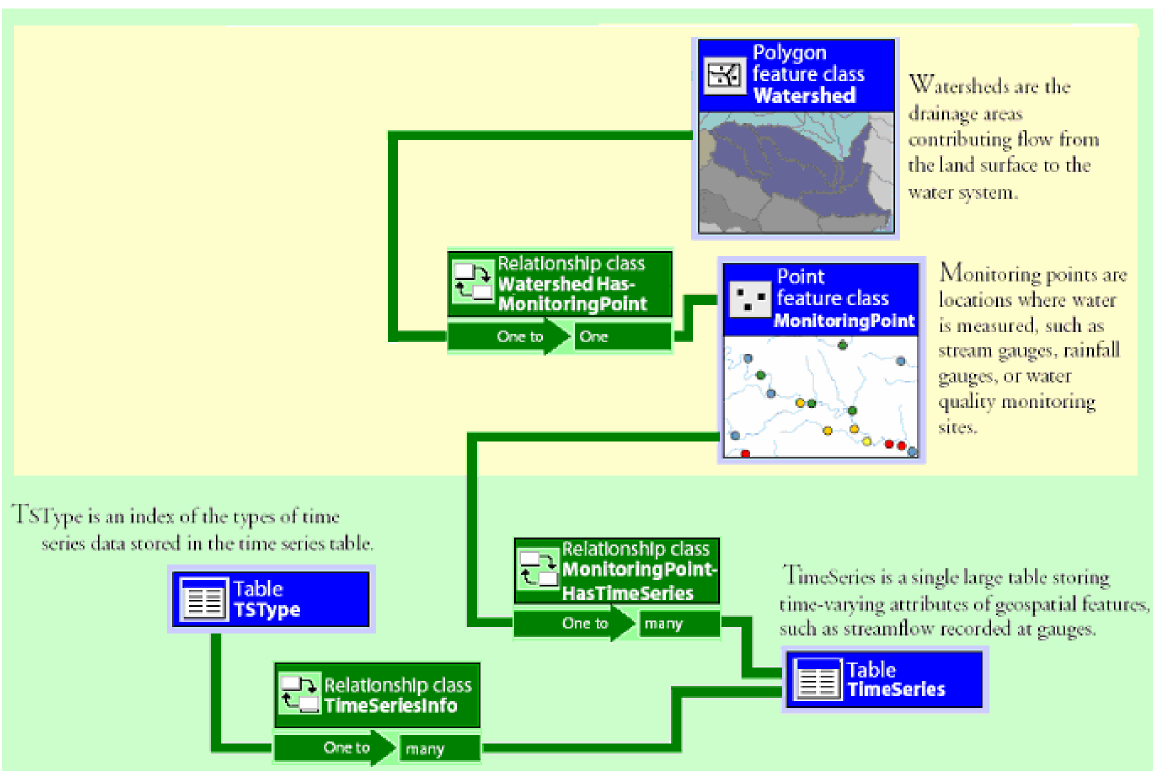


Figure 2.15 Arc Hydro Time Series relationships.

Because the Arc Hydro timeseries format contains both temporal and spatial information on every record, an Arc Hydro timeseries table can be used to manage and analyze either space-centered or time-centered data sets depending on the query used to retrieve data.

2.6 Interested Parties

2.6.1 Bexar Regional Watershed Management Coalition

The Bexar Regional Watershed Management Coalition (BRWMC) is an agreement between local agencies formed in response to the need for more coordination amongst parties interested in water resources management in the San Antonio area. The agreement was formed in 2002 between Bexar County, the City of San Antonio, and the San Antonio River Authority. Each agency has interests in water resources management from flood protection, to stomwater management, to water quality in the San Antonio River. The overarching purpose of the BRWMC is to employ a coordinated, consistent, and efficient program to manage flood control, drainage and storm water for the area around San Antonio. (SARA 2005) The BRWMC has developed a "Watershed Master Plan," which establishes goals, objectives, standards, and best management practices for water resources management in the area. It has been working on a Regional Watershed Modeling System (RWMS) with the goal of integrating water resources management in the San Antonio area.

The integration goal springs from the idea that all water resources related modeling, including hydrologic, hydraulic, and water quality attempt to simulate processing occurring in the environment. Though models for different applications simulate different water-related processes and require modeling parameters specific to the application, all the processes are occurring in the same natural environment. With this understanding, it makes sense to have the same geospatial representation of this environment in all models. To support this goal, the BRWMC is undergoing a massive data collection effort to characterize the environment in the San Antonio area and store the information in GIS format. Ongoing efforts include defining current landuse characteristics, accurately defining the locations and geometry of river channels, and redelineating watersheds with accurate topographic information.

Once data collection is completed, it will be stored in an Arc Hydro-style GIS format and be available to all interested parties in the area. The goal is to have data in the Arc Hydro database continuously updated to reflect the current state of the environment. In addition, the BRWMC hopes to use the data to develop a standard set of water quantity and quality models that can be used to evaluate future changes in the physical landscape that affect water resources. Tools under development will automatically update parameters in complex water quantity and quality models based on current conditions as reflected in the Arc Hydro database. HEC-HMS and RAS are being used as the flood models, and HSPF has been chosen for water quality modeling. (Roboyo and Maidment 2005)

2.6.2 AQUA TERRA Consultants

AQUA TERRA Consultants (ATC) is an environmental consulting firm with locations in California, Washington, Colorado and Georgia. The development for recent releases of the EPA's BASINS system as well as ongoing support for the HSPF has been the responsibility of AQUA TERRA Consultants. The BASINS system contains tools to leverage ESRI's 3.x geospatial capabilities to perform preprocessing for HSPF in a GIS environment and to support the creation of input files for HSPF. Additional tools for manipulating input data, running the model, and analyzing output data operate outside the GIS environment. The WinHSPF/GenScn/WdmUtil software package (often referred to simply as "GenScn Tools") is available free of charge from ATC's website (AQUA TERRA Consultants 2005), and provides powerful capabilities for building and calibrating HSPF models. The GenScn Tools and software were developed using the Visual Basic programming language, and designed to be extensible to provide tools for other software developers.

Tools and programming libraries developed by members of the AQUA TERRA Consultants BASINS team in Decatur, GA, are used extensively in this research. In addition, the methodology presented for preprocessing ArcGIS data for use with HSPF is closely related to the BASINS ArcView 3.x methodology.

2.6.3 CUAHSI

The Consortium of Universities for the Advancement of Hydrologic Science, Incorporated (CUAHSI) began in 2001 as a group of universities dedicated to the advancement of hydrologic science (CUAHSI 2005). The work of the Consortium is funded by a National Science Foundation (NSF) grant and has been implemented in phases over the past four years. One component of CUAHSI is the Hydrologic Information Systems (HIS) program, intended to improve the infrastructure and services for hydrologic information.

Geographic Information Systems are used extensively as a means of managing and visualizing hydrologic data in the HIS program. Because GIS provides a digital platform for an abstraction of the environment, data stored in the hydrologic information system can be easily accessed and used for a wide range of applications. Initial work of the HIS program has focused on developing a system to support the collection, management, and distribution of hydrologic information in a general way that is not tied to any specific application. Once the infrastructure for managing and distributing hydrologic data is in place, future work will be necessary to support specific mathematical modeling applications.

Chapter 3 GIS and HSPF Model Development

This chapter is intended to provide an understanding of common ways that GIS data are used to develop HSPF models. This chapter will first introduce how GIS is and is not used in HSPF modeling. Next it will present some of the common data sources and how they are used in the development of HSPF models. The HSPF model structure is very general, and GIS data is commonly used to define the specific combinations of Model Elements (or "Model Configuration") that will be used to simulate a system of interest. GIS Tools in the BASINS system are the most widely used technology for HSPF model development, and they are presented in detail. The chapter will close with a presentation of how forcing precipitation data are commonly prepared for HSPF models and how the method used for the preparation affects HSPF model configurations.

3.1 Geospatial Information and HSPF

All physically based models rely on an accurate representation of the physical environment that is to be modeled. GIS data have proven their utility in representing the hydrologic environment and GIS tools are the most widely used technology for developing input data for hydrologic models. Some of the most common tasks performed by GIS in developing models are 1) defining drainage areas, 2) calculating physically based model parameters, and 3) defining areas of the land with similar hydrologic characteristics.

HSPF and many other hydrologic and environmental models were developed long before geospatial information was widely available in a digital form. As a result, most hydrologic models, including HSPF, do not use explicit geospatial information in the model files. Rather than explicitly using geospatial information such as GIS data, model simulations are performed using physically based parameters, such as slope, area, and length that may or may not have been calculated using GIS data.

Because geospatial data is typically only used during model development, it is often not saved, maintained, or updated when changes are made to model files. Network and connectivity information must be maintained in the model files, but because there is no place to store explicit geospatial information in HSPF model files, it is often not used after initial model development. If GIS data are not updated to reflect changes to model files, inconsistencies arise between the model representation and the GIS representation of the environment used in model development. If GIS data are maintained after initial model development, they can provide an accurate representation of the areas simulated by the model and be used to facilitate the transfer of information to model files.

3.2 COMMON GIS DATA SOURCES AND HSPF MODEL PARAMETERS

The first job when setting up an HSPF model is characterizing the landscape. In recent years, GIS data for hydrologic applications has become the most widely used resource for describing the hydrologic characteristics of the land surface. Digital Elevation Models, landcover, and soil data are available for the entire US from several governmental agencies. In addition, stream networks for perennial as well as intermittent streams are available at high resolution for most areas of the country. GIS data is used widely both to estimate physically-based parameters and to define areas that have similar hydrologic characteristics. (Singh 2002)

3.2.1 Common Data Sources

The National Elevation Dataset (NED) is the most common source for DEMs and is developed and distributed by the USGS (USGS 2005). Elevation data is available for the entire US on a 30m resolution, and a large part of the country is covered with 10m resolution. Elevation data is sometimes collected on an even finer scale using a method called LIDAR (LIght Detection And Ranging). LIDAR is a remote sensing technique used to collect detailed topographic data using an aircraft mounted sensors. The NED has a vertical resolution of about 0.5 to 1 meter, however LIDAR data is capable of resolving vertical elevation changes of only about 15 centimeters. Elevation data is used to define drainage areas for hydrologic modeling. (Gueudet 2004)

The most common data source for defining the locations of river channels is the National Hydrography Dataset (NHD) (USGS and EPA 2005). The NHD is a GIS dataset containing surface water features including rivers, streams, lakes, ponds, springs, and wells. Low and medium resolution NHD data is available for the entire conterminous US, and high resolution NHD data is available for much of the country. The high resolution data contains all perennial as well as most intermittent streams, drainage ditches, and other surface water features. The NHD data has recently been converted to a form called NHDinGeo, which is distributed in ESRI Geodatabase format and includes network-tracing capabilities. Lines through water bodies are inferred so that a continuous stream network is available for the entire US. NHD data is typically used to define river segments to be simulated by HSPF models. The use of NHD and the process of developing digital descriptions of river networks for hydrologic modeling is not directly addressed in this research.

A major factor affecting the hydrologic behavior of an area of land is vegetative cover. Vegetation has a large influence on the movement of water over the land surface, into groundwater, and out of groundwater through evapotranspiration. The vegetative cover of an area of land is typically inferred from GIS data describing the land cover or landuse characteristics. Landuse data is probably the most common information used to characterize the landscape for HSPF modeling. Since vegetation affects many hydrologic processes, it is often assumed that areas with similar landuse characteristics can be simulated with the same or similar Operations in HSPF models.

The Multi-Resolution Land Characteristics (MRLC) Consortium is a group of federal agencies, including the USGS (United States Geological Survey), EPA and

NOAA (National Oceanic and Atmospheric Administration), that has been working to maintain a nationally consistent dataset of satellite and remote sensing landuse Cover data. In 1993, the Consortium purchased satellite imagery dating between 1989 and 1992 and developed a dataset called the National Land Cover Dataset (NLCD) 1992 for the entire conterminous United States. A recent effort beginning in 1999 is underway to purchase newer satellite imagery and develop a NLCD 2001. (EPA, USGS, and NOAA)

Another common Land Cover dataset is the Land Use / Land Cover (LULC) dataset which was developed by the USGS in the early 1990's and is based primarily on aerial photography from the 1970's and 1980's. The LULC data, commonly referred to as GIRAS landuse data, is distributed mostly in the form of shapefiles (vector GIS data) while the NLCD data is typically distributed in a raster form. (EPA 2005d)

Soil Type is another characteristic commonly used GIS dataset for describing the landscape. The Natural Resource Conservation Service (NRCS), formerly the Soil Conservation Service (SCS), is responsible for collecting and distributing soil survey information for the United States. The two NRCS datasets most widely used for natural resource planning and management are the Soil Survey Geographic Database (SSURGO) and the State Soil Geographic Database (STATSGO). The SSURGO database is the most detailed database available and is intended for use on a county/township/landowner scale. The STATSGO database is a generalization of the SSURGO database and is intended for regional/multistate/river basin scale planning and management (NRCS 2005a, NRCS 2005b)

3.2.2 GIS Data and Model Parameters

In HSPF, Operation Modules simulate hydrologic and water quality processes using Operation Parameters. These parameters are used to represent the characteristics of the landscape. A complex HSPF model will use hundreds or even thousands of parameters in the equations used to simulate the movement of water through the hydrologic environment. Some of these parameters represent physically based characteristics such as slopes, lengths, and areas and can be directly estimated or calculated from readily available GIS data. While GIS data is widely available for some parameters, other characteristics such as surface roughness and infiltration capacity cannot be easily inferred or measured from GIS data. Though these parameters cannot be calculated directly, they are often assumed to be uniform over areas with similar landuse characteristics. Areas of the land that should have similar characteristics are typically inferred from information about soils, landuse, and other GIS data. For instance, it is often assumed that all forested land will have similar interception storage capacity, and areas with similar soil types will have similar soil moisture capacities. These areas, with common hydrologic characteristics, define the areas that will be simulated with individual HSPF Operations.

Table 3.1 shows a list of HSPF model parameters, their associated characteristics, and the type of data that is typically used to either calculate them, or to define areas with uniform characteristics. This research is chiefly concerned with developing input data for hydrology simulations on Land Segments in HSPF models, and only the parameters for the simulation of water movement over the land surface are presented. Some of the physically based parameters can be calculated readily from widely available GIS data, while others are used as calibration parameters. The table gives the data source that is typically used to either calculate the parameter, or define areas with uniform characteristics relevant to the parameter. Further information on details of HSPF model parameters can be found in “BASINS Technical Note 6” (EPA 2000).

Table 3.1 HSPF model parameters and common data sources.

Parameter	Description	Type of Data	Data Source
FOREST	Percent Forested	Land Cover data	LULC / NLCD
LZSN	Lower zone nominal moisture storage in soil	Calibration	
INFILT	Infiltration parameter for soil	Calibration	
LSUR	Length of overland flow plane	Digital Elevation Model	NED / LIDAR
SLSUR	Slope of overland flow plane	Digital Elevation Model	NED / LIDAR
KVARY	Parameter for active groundwater	Land Cover and/or Soil data	LULC / NLCD - SSURGO / STATSGO
AGWRC	Active groundwater recession coefficient	Land Cover and/or Soil data	LULC / NLCD - SSURGO / STATSGO
PETMAX	Temperature for maximum evapotranspiration	Land Cover data	
PETMIN	Temperature for minimum evapotranspiration	Land Cover data	
INFEXP	Infiltration exponent parameter	Land Cover and/or Soil data	LULC / NLCD - SSURGO / STATSGO
INFILD	Parameter for variability of infiltration capacity	Land Cover and/or Soil data	LULC / NLCD - SSURGO / STATSGO
DEEPFR	Parameter for sending water to inactive groundwater	Calibration	
BASETP	Parameter for evaporation from base groundwater	Calibration	
AGWETP	Parameter for evaporation from activegroundwater	Calibration	
CEPSC	Interception storage capacity	Land Cover data	LULC / NLCD
UZSN	Upper zone nominal moisture storage in soil	Land Cover and/or Soil data	LULC / NLCD - SSURGO / STATSGO
NSUR	Surface roughness	Land Cover and/or Soil data	LULC / NLCD - SSURGO / STATSGO
INTFW	Interflow parameter	Calibration	
IRC	Interflow recession coefficient	Calibration	
LZETP	Parameter for evaporation from lower soil zone	Calibration	

The only truly physically based parameters shown in Table 3.1 are LSUR (length of overland flow plane) and SLSUR (slope of overland flow plane). Some of the other parameters do have physical meaning, but are not used in completely physically-based equations. The UZSN (upper zone nominal storage) is intended to be representative of the residual moisture content of the soil, however, it is often used as a calibration parameter in the somewhat empirical equations used to simulate infiltration. GIS data are used to define the areas of land (Model Elements) that will be assumed to have a common set of parameters. However, the values for these parameters are often changed during the model calibration process and not directly measured or inferred from GIS data.

3.3 HSPF MODEL CONFIGURATIONS

HSPF models simulate a system of interest using a network of HSPF Operations which each simulate a finite portion of the real world. To simulate a hydrologic system, HSPF was designed to have two types of Operations to simulate the land surface, Pervious Land Segments (PERLND) and Impervious Land Segments (IMPLND). These Land Segment Operations are typically linked to Reach Segment Operations (RCHRES), which simulate the flow of water in rivers.

Land Segments flowing to Reach Segments is inherent to the movement of water in the environment and the obvious choice for structuring an HSPF model. The choice of how many Land and River Segments to be use, and how they are linked together, however, is made by the modeler. The following sections will discuss different HSPF "model configurations," or methods for linking Land and River Segments together.

3.3.1 Motivating Factors for Model Configurations

One motivating factor for configuring an HSPF model is the desire to capture the spatial variability of processes occurring on the land surface. If different areas of a watershed contain similar land surface characteristics, it is likely that the hydrologic processes occurring on these different areas will be similar. HSPF Operations make use of this assumption by simulating a vertical water balance on one or more Land Segments land and distributing the results to the appropriate River Segment. Landuse or soils data are typically used to characterize the hydrologic characteristics of the land surface. To capture and simulate the variability of hydrologic processes on the land surface, different HSPF Land Segment Operations are used (with different relevant parameters) to simulate separate parts of the watershed.

Another motivating factor in the configuration process is the objective of the model. The objective of a hydrologic or water quality model may require that results be produced at specific points in a river network, for instance, identifying the peak flow in a river near a populated area, or the concentration of bacteria near a water intake. These objectives typically require that a river segment be divided at a location where a monitoring station exists so that model results can be directly compared with observed data. Additionally, River Segments must be divided at the junctions where two or more rivers or streams converge to a single stream. In addition to objectives concerning results in the river network, hydrologic models are also used to investigate the effects of changes to the land surface. The configuration of Land Segments in an HSPF model can be designed to examine a specific change to the land surface by adjusting the parameters of individual Land Segment Operations.

The availability of forcing data, especially rainfall, is also an important factor for model configuration. The structure of an HSPF Operation, for instance a Land Segment, requires that Forcing Data such as rainfall be distributed uniformly over an Operation. HSPF Operations are often created to represent an area with uniform land surface characteristics, but because of the structure of the HSPF model, the entire area represented by an Operation must also receive uniform rainfall. Though two areas of land may have identical hydrologic characteristics on the land surface, if they do not receive the same amount of rainfall, they must be modeled with two separate HSPF Operations. The way in which precipitation data affects HSPF model configurations is presented in Section 3.6.

If the area to be modeled with HSPF can be assumed to receive uniform precipitation, the distribution of rainfall need not be considered when developing the configuration of a model. However, if there is more than one rain gauge in the area, or if other distributed rainfall information (such as radar estimates of rainfall) is to be used, the spatial distribution of precipitation must be considered when configuring a model. In this case, two areas of land that have uniform surface characteristics but receive different rainfall data are represented by separate HSPF Operations.

Sometimes, River Segments are divided into smaller segments because of significant changes in slope or river – aquifer interactions. In HSPF, the flow of water through River Segments is modeled using a simple volume or stage vs. discharge relationship. A lumped flow routing scheme is applied using an invariable, single valued storage function relating discharge from the segment to storage in the segment (Bicknell 2001). This research is chiefly concerned with segmenting the land surface for HSPF modeling, but the model configuration of a river system is related to the chosen segmentation for the land surface.

Before presenting a specific method for PreProcessing GIS data to define HSPF model configurations, the following section will present increasingly complex methods for dividing the landscape into HSPF model Operations and and discuss the implications they have on the assumptions of the resulting HSPF model.

3.3.2 HSPF Model Configuration Examples

An HSPF model could be set up in a completely lumped[1] manner in which only a single Pervious Land Segment is simulated to contribute water and other constituents to a

---

1 The terms "Lumped" and "Distributed" have specific implications when describing mathematical models. These implications have to do with not only the spatial scale of a model representation, but also the techniques of solution. Strictly speaking, HSPF and Operations within the model are Lumped models because they do not consider the partial derivatives of processes with respect to space in simulations. (Singh 2002) Despite these implications, in this paper when describing HSPF model configurations, the terms "Lumped" and "Distributed" will be used to describe the degree to which the land surface and river systems have been broken up into smaller pieces. "Lumped" will mean that large areas of land which may or may not have similar spatially variable properties are simulated as a single Land Segment. "Distributed" will mean that an attempt is made to segment the model so that spatially variable parameters and data are relatively uniform over an individual Land Segment.

single River Segment as shown in Figure 3.1. In Figures 3.1, 3.2 and 3.3 (as well as 3.14 and 3.15) the left (GIS) figure illustrates the “real world” and the right (schematic) figure shows the HSPF model representation. This completely lumped configuration would be appropriate for only the simplest case, if the land surface characteristics did not vary widely over the watershed, the watershed received uniform rainfall, and if there was no need for information at points internal to the watershed or stream network.

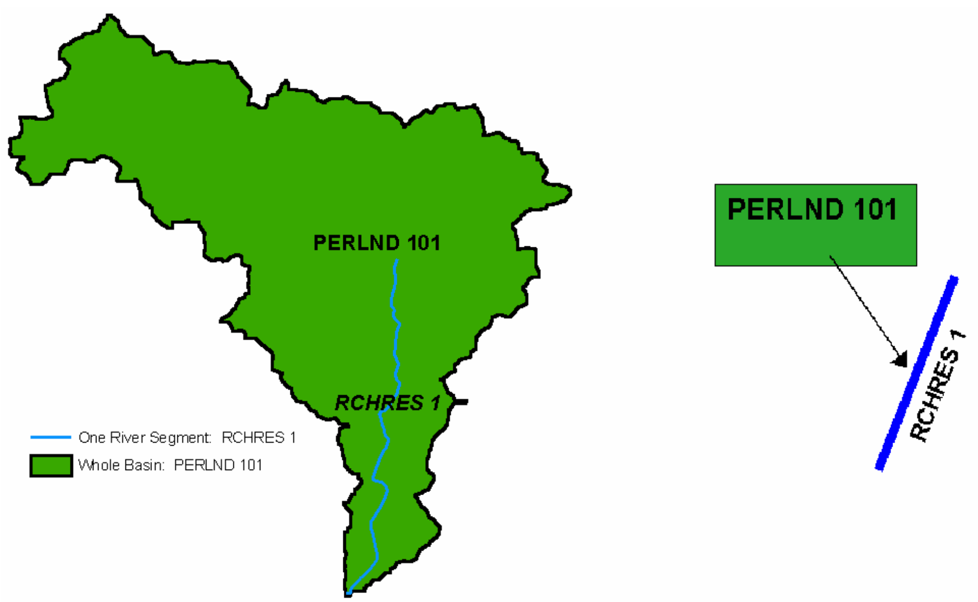


Figure 3.1 Completely lumped Land Segment configuration – Single drainage area.

If the objectives of the modeler require that output be calculated somewhere in the middle of the river segment, a lumped representation will not suffice and multiple RCHRES Operations are required. If the land surface characteristics do not vary over the watershed and uniform rainfall is assumed, the single Land Segment could still be used, but it would contribute to each river segment in proportion to its drainage area as shown in Figure 3.2.

In the configurations shown in Figures 3.1 and 3.2, the entire land area is simulated using a single HSPF Land Segment Operation. A single set of parameters assigned to the Land Segment Operation, and the results of hydrologic processes simulated on the Land Segment are assumed to be characteristic of the entire land area. An HSPF Land Segment Operation essentially calculates water fluxes (input from atmosphere, surface runoff, infiltration, etc.) on an areal basis and distributes the output to river segments in proportion to drainage area. The results of the Land Segment simulation will be distributed to each Reach Segment shown in Figure 3.2 in proportion to its drainage area.

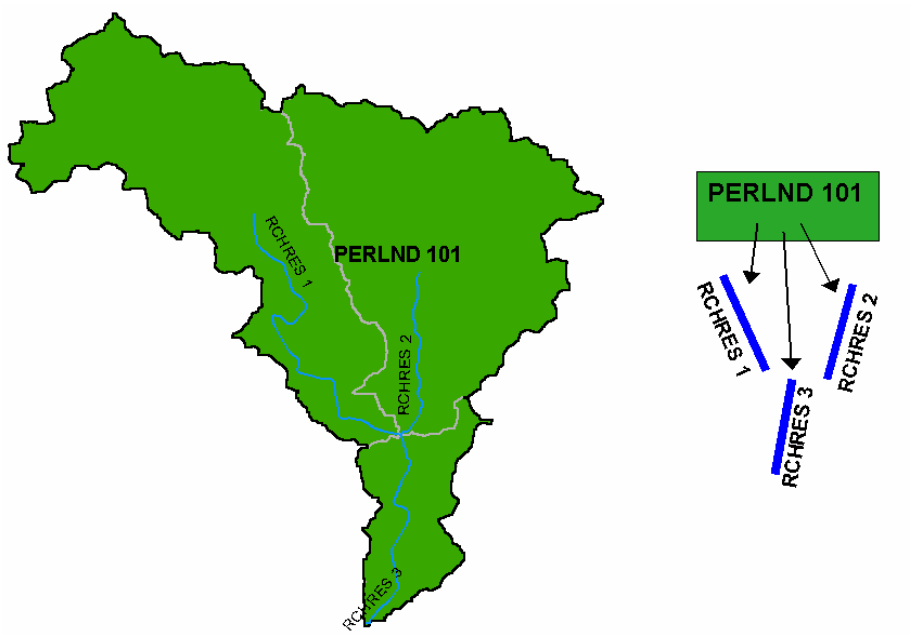


Figure 3.2 Completely lumped Land Segment configuration – Multiple drainage areas.

In order to capture the spatial variability of hydrologic processes on the land surface, most HSPF modeling applications further subdivide the land surface based on land surface characteristics such as landuse, slope, or soil type. Figure 3.3 shows an example of an HSPF configuration in which five different types of Land Segment Operations contribute to three Reach Segments. In Figure 3.3, the Land Segments are defined according to Land Cover characteristics, but the landscape could be separated according to soil type, slope, or any other data that is available.

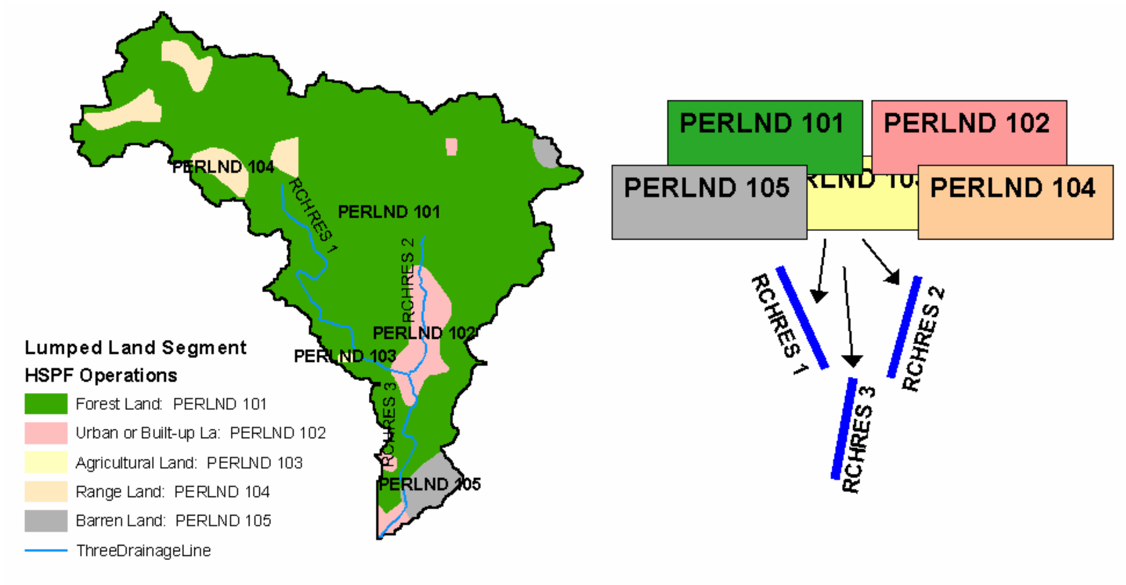


Figure 3.3 Distributed Land Segment configuration – Single set of Land Segments.

Figure 3.3 shows several separate areas designated as range land, however, the schematic (which shows the HSPF model representation) shows only one PERLND (104). The rationale behind this choice of representing spatially incontiguous areas of land with a single Land Segment is related to the method used to simulate water movement over the land surface. HSPF Model Operations essentially compute a vertical water balance, and if each of the areas shown as "Range Land" in Figure 3.3 has similar land surface characteristics, there is no reason for believing that an areal water balance will be different for each one. In this case they could all be simulated using a single HSPF Operation even though they are not spatially contiguous.

If separate areas of land are lumped together into a single Operation, some of the physically-based parameters for the range land Operation (such as overland flow distance and slope) may not be physically-based. However, if average values can be used to adequately characterize the process, these parameters can be assumed to be characteristic of all the range land areas. If some areas of Range Land reside in drainage areas for different River Reaches, the Output from the single Range Land Operation must be weighted according to how much resides in each drainage area and dispersed to the appropriate River Segment. A single Land Segment Operation can contribute flow to one or more River Segment Operation.

The configurations shown in Figures 3.1 to 3.3 explicitly account for some spatial variability in land surface hydrologic processes by using different model operations to simulate areas of land with different land surface characteristics. However because an Operation must receive uniform forcing data, in each configuration presented above, uniform rainfall must be assumed over the entire model area. This type of configuration is appropriate if the area is small enough to assume uniform rainfall or if limited rainfall information is available. Even if rainfall is not truly uniform over the entire simulated area, rainfall data from a single gauge is often used because it represents the best information available.

If rain gauge information is available at more than one point in or near the watershed, Thiessen polygons could be used to assign rainfall information to different areas of the watershed. Figure 3.4 shows Thiessen polygons surrounding three rain gauges in the proximity of the watershed. Many HSPF models assume that these Thiessen areas each receive uniform rainfall corresponding to the nearest rain gauge. Most HSPF modeling applications employ landuse data to initially divide the land into different Operations, and further subdivide these areas to receive rainfall from the appropriate rain gauge.

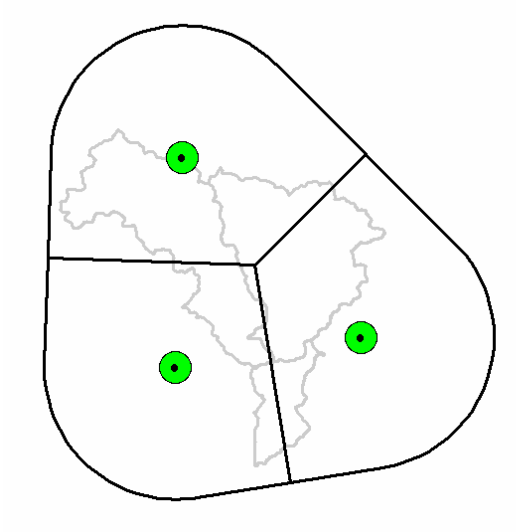


Figure 3.4 Thiessen rain areas.

Figure 3.5 shows a configuration that is capable of using the rainfall information from the rain gauges in Figure 3.4. While the configuration from Figure 3.3 must assume uniform rainfall over the entire land area, the configuration in Figure 3.5 uses separate Operations to simulate the response of each landuse type within different Thiessen rain gauge areas.

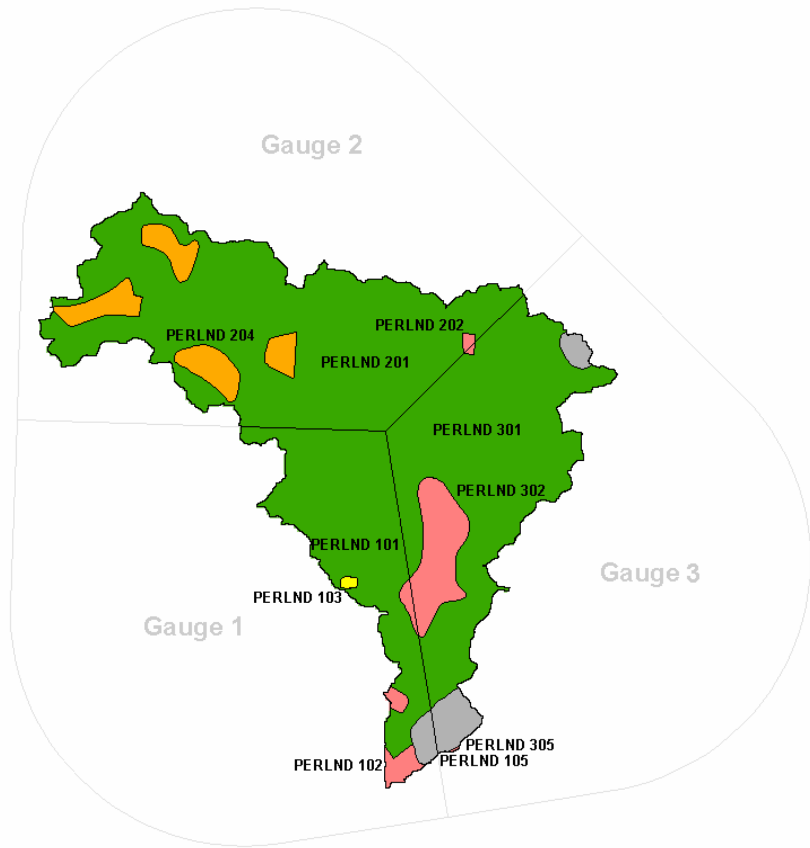


Figure 3.5 Model configuration for Thiessen rain gauge precipitation.

Though many or all of the parameters related to surface characteristics may be the same across Thiessen rain gauge areas, the separation is necessary to allow for the use of the more distributed rainfall information. In this configuration, the land is divided into categories based on surface characteristics and a unique set of these categories is used to receive forcing data (precipitation and evaporation) from each of the Thiessen rain gauge areas.

It is almost certain that some of the unique combinations of Thiessen gauge and Land Cover type (each model Operation) will fall within the drainage area of more than one River Segment. Output from the Land Segments must be distributed to the appropriate River Segments in proportion to the amount that lies within each drainage area. It is apparent from the above presentation that the configuration of a large HSPF model with multiple River Segments, Land Cover types, and rain gauges would be extremely complicated. GIS data can be used to overlay many large data sets to define each of the areas described above. Catchment boundaries can be combined with data describing areas having uniform land surface characteristics and areas receiving uniform rainfall to efficiently and accurately define a complex model configuration such as that presented in Figure 3.3 and 3.5.

3.4 Operation Numbering Conventions

The Operation Number is a number used internally by the HSPF model software to identify each Operation to be simulated. The Operation Number is unique within an Operation Type (PERLND, IMPLND, and RCHRES), but Operations from different types may share the same Operation Number. For example, PERLND 101, IMPLMD 101, and RCHRES 101 could all exist within the same HSPF model. The HSPF model uses the Operation Number in combination with the Operation Type to uniquely identify each Operation within the model. There are no rules built into the HSPF model for numbering Operations, except that they must not exceed three digits. They do not need to begin with one or be in sequence or follow any pattern at all. If an HSPF model has five PERLND Operations, they can be numbered in any way: [1, 2, 3, 4, 5] or [101, 102, 103, 104, 105], or [461, 53, 115, 879, 7]. While any of these numbering systems would work, a logical methodology for numbering operations is extremely helpful for working with a large, complex HSPF model.

Though no spatial information is explicitly stored in the .uci file, often the Operation Number is assigned in a way that implicitly stores some spatial information. As an example, the configuration shown in Figure 3.5 uses a logical system to aid in organization. The first digit corresponds to the Thiessen rain gauge area and the last digit corresponds to the Land Cover type.

While such a convention is not necessary, it is extremely useful when editing a complex model. For the convention adopted above, it would be easy to find all Model Operations with the same landuse by looking at the last digit of the Operation Number. When the model is being calibrated and parameters relating to landuse characteristics are being changed, care can be taken to assign the same parameters to the Operations with similar landuse.

Using the Operation Number to store geospatial information has limitations because of the three-digit limit for HSPF Operation Numbers. For the convention adopted in Figure 3.5, there could not be more than 10 types of land simulated by the model because (in the decimal system) the final digit of the Operation Number is limited to 0-9. Additionally, no more than 100 thiessen rain areas could be used because the first two digits of the Operation Number is limited to 0-99. For most HSPF models ten landuse types is sufficient, and the necessity of a model with more than 99 thiessen areas is highly unlikely. However, it is apparent that using the Operation Number to store spatial information may not be practical for some applications.

Other limitations to the numbering conventions presented above arise when it becomes necessary to add additional segments to a model or change a model in a way that is not consistent with the Operation Numbering convention. For instance, after a model has been developed using the convention presented above, rainfall data may later become available from additional rain gauges, or land surface characteristics may change. If new rainfall data is to be used, and new land surface characteristics are to be modeled, it will be necessary to divide, modify, or add additional Land Segments in the model. The convention adopted for numbering Operations may not be structured in a way that allows for these changes while maintaining a consistent link between Land Segment Operation Numbers and the land surface they are intended to simulate.

The Operation numbering convention which is used by the BASINS preprocessing system and the one that will be used in this research uses only landuse and uniform – rain areas to define Land Segments for HSPF modeling. Though any number of Land and River Segment Operations could theoretically be linked together and simulated with HSPF, Version 12 of the program is limited to 500 Operations for a single run.

For configurations such as those presented in Figure 3.3 and 3.5, 'Lumped' land segments are simulated and contribute to the appropriate River Segment in proportion to their drainage area. In this type of configuration, the 500 Operation limit is rarely exceeded. However, Section 3.6 presents a more 'Distributed' Land Segment configuration, adopted to make use of distributed precipitation data. In this configuration, a unique set of Land Segments is used to simulate the drainage area for each River Segment. With up to 10-15 Land Segments used to simulate the drainage area for each River Segment, the 500 Operation limit could easily be exceeded.

3.5 BASINS HSPF PREPROCESSING METHODOLOGY

Chapter 2 presented many applications that have been developed with the purpose of preprocessing GIS data for Hydrologic modeling. The most popular application with the relevance to HSPF is the BASINS system. The BASINS system contains GIS tools to aid in the development of HSPF models, and non-GIS tools to work with HSPF model files after initial development (EPA 2001). This section will outline the methodology implemented by the BASINS system to create a new HSPF model starting with GIS data. The BASINS HSPF Preprocessing methodology is the most widely used method for creating new HSPF models and provides the basis for the development of an ArcGIS HSPF Preprocessing methodology (presented in Chapter 4).

3.5.1 Overview

The GIS component of the BASINS 3.x system operates within the ArcView 3.x environment, a proprietary software package from ESRI. The GIS environment contains extensive tools for gathering, organizing, and summarizing data, and several tools are included to aid in preprocessing data for HSPF models. GIS tools automatically perform the spatial analysis required to define HSPF model configurations such as that shown in Figures 3.3 and 3.5. Though an HSPF model can be configured based on any number of different land surface characteristics, the BASINS tools for automatically creating an HSPF model are designed to use landuse as a means of dividing the landscape.

After performing some basic tasks such as delineating watershed boundaries and importing landuse and soils data using standard BASINS tools, an HSPF modeler has the option to "Create a new HSPF project" as depicted in Figure 3.6. This option sets into motion a set of screens that guides the user through selecting a name for the new HSPF project (.uci file), selecting the type of landuse data to use, and defining the amount of impervious land in each landuse category. These screens are illustrated in Figures 3.7.

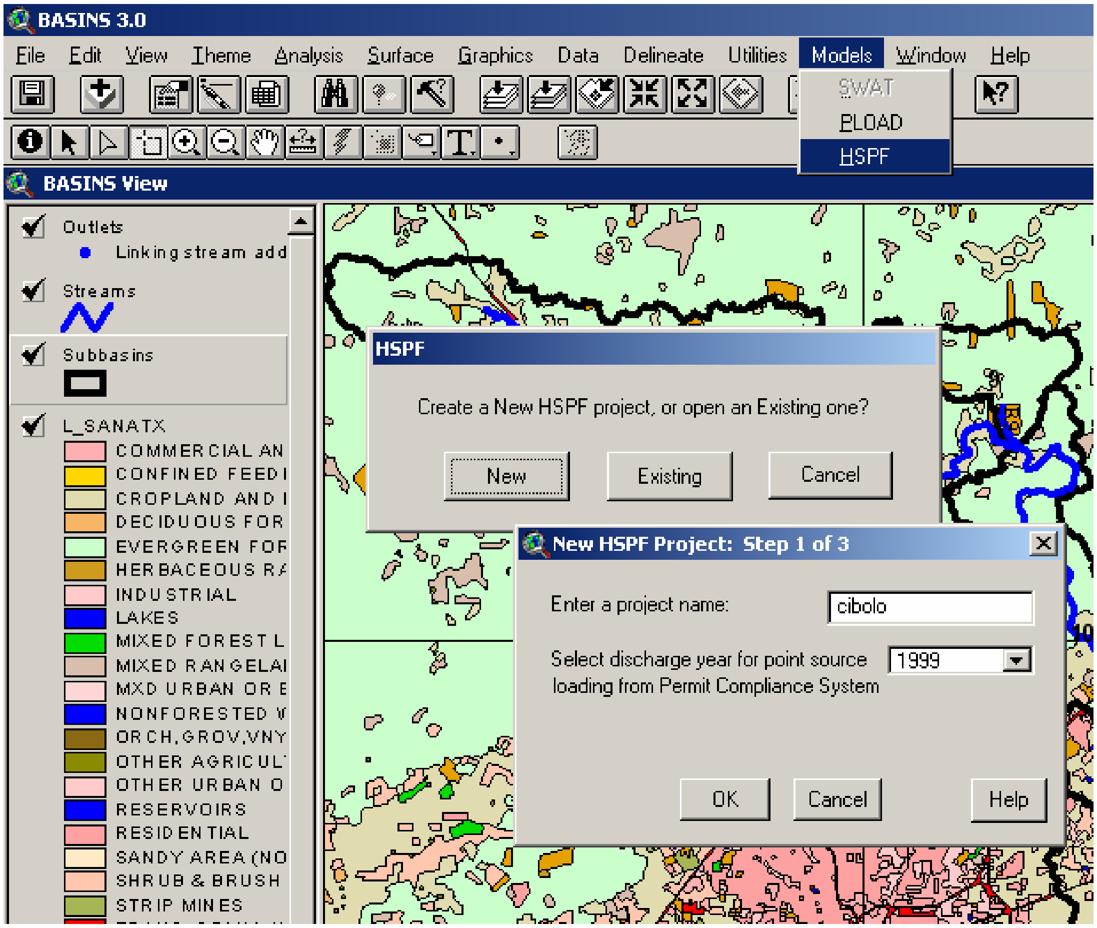


Figure 3.6 BASINS HSPF tools: Access from ArcView 3.x.

At this point, control is passed to another program outside of the GIS environment, but still part of the BASINS system, WinHSPF. An additional screen, shown in Figure 3.7, prompts the user for meteorological data to use in driving the HSPF model, and output files for HSPF results. Figure 3.7, shows an option for “Model Segmentation.” This involves either creating a “Grouped” or “Individual” model, which defines the type of model configuration that will be used in building the new .uci file.

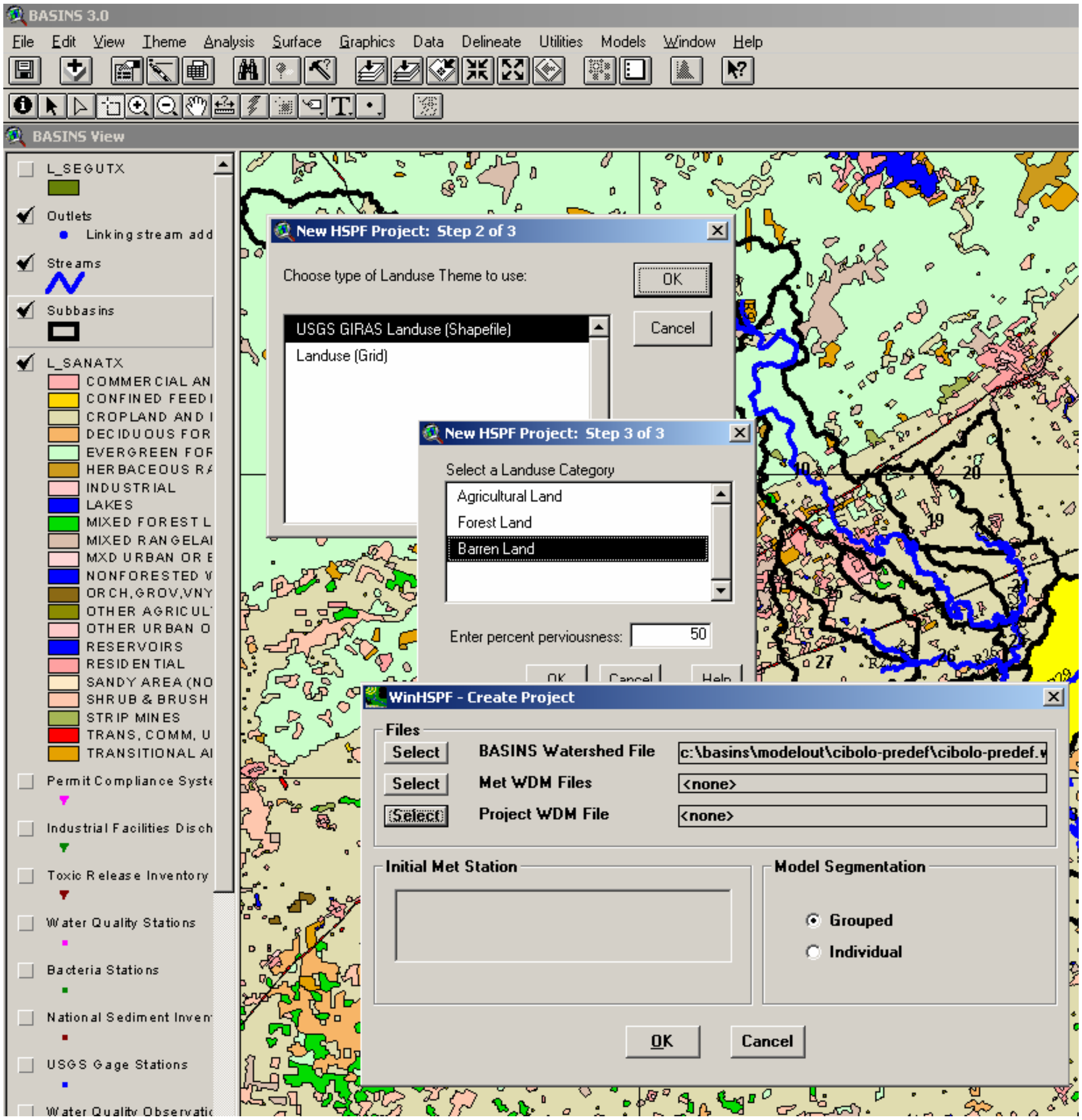


Figure 3.7 BASINS HSPF tools: Define landuse, choose configuration.

In the “Grouped” option, only one Land Segment is created for each landuse category. Each “Grouped” Land Segment (PERLND Model Operation) will have constant parameters and receive uniform Forcing Data, but contribute to River Segments in proportion to the appropriate drainage area. In the “Individual” option, a set of Land Segments (one for each category) is created for each River Reach Operation. Each set of “Individual” Land Segments contributes to only one River Segment, and can receive its own input Forcing Data. The implications of these different configurations will be discussed further in Section 3.6.

After completing the BASINS HSPF Preprocessing tools, a new .uci file (and thus HSPF model) is created and opened for editing within the WinHSPF environment. The entire process takes only a few minutes and allows users who are relatively unfamiliar with the intricacies of the HSPF model structure to create a simple, but functional model. The details of the procedures used to calculate the geospatially related parameters for HSPF Operations and the complexities involved in managing and transferring data are hidden behind the relatively simple user interface. While the simplicity is appreciated by many users of HSPF, the curious may wonder what is going on behind the scenes.

3.5.2 Parameters from Shapefiles

If the GIS tools available in BASINS are used for delineating watersheds, three GIS shapefiles are produced to represent the Subbasins, Streams, and Outlets of each catchment. The BASINS HSPF Preprocessing tools were developed to work with these three shapefiles as created within the BASINS system, however, shapefiles developed with external tools can be used if they contain the necessary information stored in an appropriate format. These three shapefiles as well as landuse data must be present to use the BASINS HSPF Preprocessing tools. Table 3.2 summarizes the attributes that are calculated using BASINS watershed delineation tools and, consequently, the ones that must be present when using the BASINS "Predefined Delineation" by importing shapefiles from another source.

Table 3.2 Attributes of the HSPF-related BASINS GIS data.

Shapefile	Field Name	Description	Necessary for BASINS tools	Source
Streams	Subbasin	Subbasin number	***Required***	Stream network
Streams	Subbasinr	Downstream Subbasin	***Required***	Stream network
Streams	Numin	Number of inlet subbasins	Not Used	
Streams	Areac	Cumulative drainage area [hectares]	Not Used	
Streams	Len2	Stream reach length [meters]	***Required***	Shape_Length
Streams	Slo2	Stream reach slope [%]	***Required***	Calculated from MinEl, MaxEl, Length
Streams	Wid2	Stream reach width [meters]	***Required***	Detailed Cross Section Information
Streams	Dep2	Stream reach depth [meters]	***Required***	Detailed Cross Section Information
Streams	MinEl	Minimum elevation of the stream reach [meters]	***Required***	Streams and DEM
Streams	MaxEl	Maximum elevation of the stream reach [meters]	***Required***	Streams and DEM
Streams	Sname	String available for labeling the theme	Used (but not required)	Name for Stream
Subbasins	Subbasin	Subbasin number	***Required***	Stream - Subbasin Connection
Subbasins	Area	Subbasin area [hectares]	Not Used	
Subbasins	Len1	Stream Reach (longest path within the subbasin) length [meters]	Repeated / Not Used	
Subbasins	Slo1	Subbasin slope [%]	***Required***	Subbasin / DEM or Slope Grid
Subbasins	Sll	Field slope length [meters]	Not Used	
Subbasins	Csl	Reach (longest path) slope [%]	Not Used	
Subbasins	Wid1	Reach width [meters]	Repeated / Not Used	
Subbasins	Dep2	Reach depth [meters]	Repeated / Not Used	
Subbasins	Latitude	Latitude of the subbasin centroid [decimal degrees]	Not Used	
Subbasins	Elevation	Elevation of the subbasin centroid [meters]	Not Used	
Subbasins	Bname	String available for labeling the theme	Not Used	
Outlets	Xpr	X coordinate in the current projection	only if point source	
Outlets	Ypr	Y coordinate in the current projection	only if point source	
Outlets	Lat	Latitude [decimal degrees]	only if point source	
Outlets	Lon	Longitude [decimal degrees]	only if point source	
Outlets	Type	Outlet type	only if point source	
Outlets	ID	Outlet ID	only if point source	
Outlets	Pcsid	Unique ID from respective program system	only if point source	

(green)	*physically based parameter calculated with preprocessing tools*
(light green)	*physically based parameter calculated with other tools*
(light blue)	*required parameter used to define stream and subbasin connectivity*
(yellow)	*name parameter used for streams, but not required*

Though all of these attributes must be present in the shapefiles for the BASINS HSPF Preprocessing tools to function correctly, some of them are redundant and many are not used during HSPF preprocessing. In addition to having repeated information, such as the “Stream reach length” being stored on both the Streams shapefile and the Subbasins Shapefile, many of the fields are not explicitly used in building a new HSPF model. The details of which attributes are required for HSFP model development and which are used for other BASINS functionality are color coded in Table 3.2. Twelve of the parameters are required for the development of HSPF models, and seven of those are physically-based and require the use of GIS calculations.

3.5.3 Defining Model Configuration

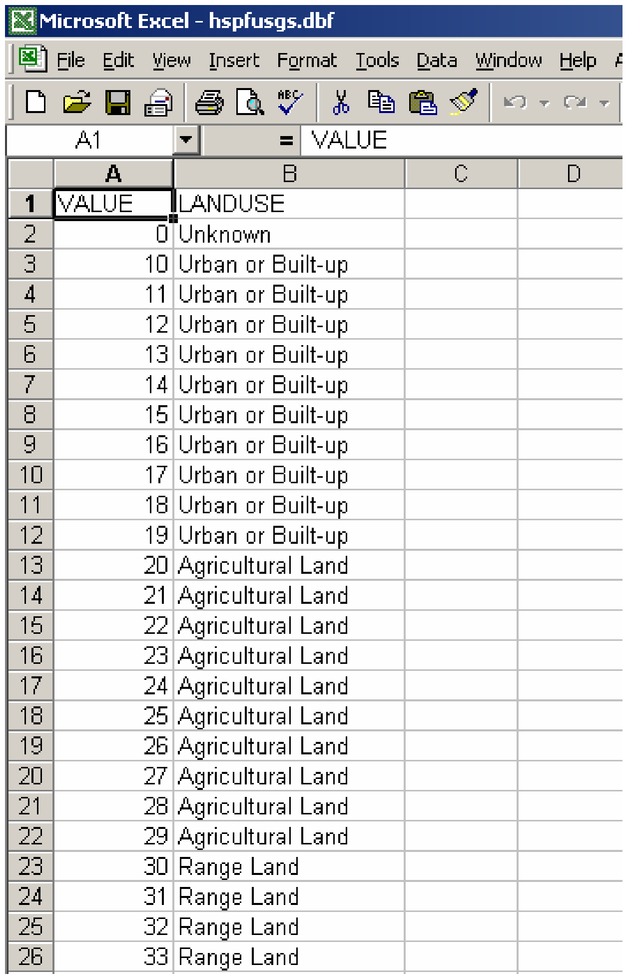
The BASINS HSPF Preprocessing tools divide the landscape according to landuse categories as shown in Figure 3.3. They use landuse data in conjunction with the Subbasins shapefile to calculate how much area of each landuse category drains to each river segment. Often, GIS landuse data contains more than 30 different categories, which is more detail than a typical HSPF model requires.

For instance, the Land Use Land Cover (LULC) GIRAS dataset is a popular source for GIS landuse data and it contains one digit (5-10 categories) and two digit (30-50 categories) classifications[2]. GIS landuse data often contains a two-digit classification code. However, the modeler may want to only use the 1 digit classification (~5-10 categories) when configuring his model. Alternatively, an HSPF user may be modeling an intensely agricultural area and want to have detailed classifications for different types of agricultural land.

The BASINS HSPF Preprocessing tools use information from an additional table, called “hspfusgs.dbf,” to define landuse categories more relevant to HSPF modeling.

[2] BASINS also has an option for using raster landuse data instead of the polygon GIRAS dataset, but only the vector option will be presented here.

The hspfusgs.dbf table is used to map landuse categories from GIS data to landuse types to be simulated by HSPF. Figure 3.8 shows the hspfusgs.dbf table mapping two-digit LULC categories to one digit types for use in HSPF modeling.



	A	B	C	D
1	VALUE	LANDUSE		
2	0	Unknown		
3	10	Urban or Built-up		
4	11	Urban or Built-up		
5	12	Urban or Built-up		
6	13	Urban or Built-up		
7	14	Urban or Built-up		
8	15	Urban or Built-up		
9	16	Urban or Built-up		
10	17	Urban or Built-up		
11	18	Urban or Built-up		
12	19	Urban or Built-up		
13	20	Agricultural Land		
14	21	Agricultural Land		
15	22	Agricultural Land		
16	23	Agricultural Land		
17	24	Agricultural Land		
18	25	Agricultural Land		
19	26	Agricultural Land		
20	27	Agricultural Land		
21	28	Agricultural Land		
22	29	Agricultural Land		
23	30	Range Land		
24	31	Range Land		
25	32	Range Land		
26	33	Range Land		

Figure 3.8 hspfusgs.dbf table, used to define landuse categories.

The hspfusgs.dbf table resides in a folder within the BASINS working directory on the computer. This table must be edited manually or using tools in the BASINS system in order to change the types of landuse to be simulated by a new HSPF model.

3.5.4 Default Parameters and Intermediate Text Files

Some physically based parameters are stored in the Subbasins and Streams shapefiles, however, many parameters are not readily available from GIS data. These parameters are still needed to create a new HSPF model even if only default values are assumed. The BASINS HSPF Preprocessing tools use default values stored in a “starter .uci file,” (starter.uci) and an “HSPFMsg database” (hspfMsg.mdb) to assign parameters that are not available from GIS data.

Also hidden behind the scenes of the BASINS HSPF Preprocessing tools are intermediate files used to extract necessary information from the GIS data. Four text files, .rch, .wsd, .ptf, and, .psr are created by the GIS preprocessing tools and contain information from the GIS data necessary to build a new .uci file.

- .wsd File – used to create Land Segment Operations
- .rch File – used to create Reach Segment Operations
- .ptf File – used in defining the Channel Geometry (FTABLES)
- .psr File – only used if including point sources from BASINS

The .wsd file stores information used to define the Land Segments to be simulated by HSPF. The .rch file contains information defining the connectivity of the river network and other parameters to define River Segments to be simulated by HSPF. The .ptf file defines the channel geometry and is used to build the FTABLES Block of text lines in the .uci file. The .psr file is used only if point sources of pollutant loads are being included in the initial .uci file. Examples of these files as well as a detailed description of the content and format are given in the WinHSPF user’s manual available from the EPA’s BASINS website. The data in these intermediate text files are combined with information from the “starter.uci” and the “HSPFMsg database” to create a new .uci file.

The final step in the BASINS HSPF Preprocessing methodology is building a new .uci file. Though the BASINS tool calls it from the GIS program interface, an external program, WinHSPF, is actually used to create a new .uci file from the information stored in the intermediate text files.

3.5.5 Summary

Figure 3.9 illustrates the components of the BASINS HSPF Preprocessing methodology.

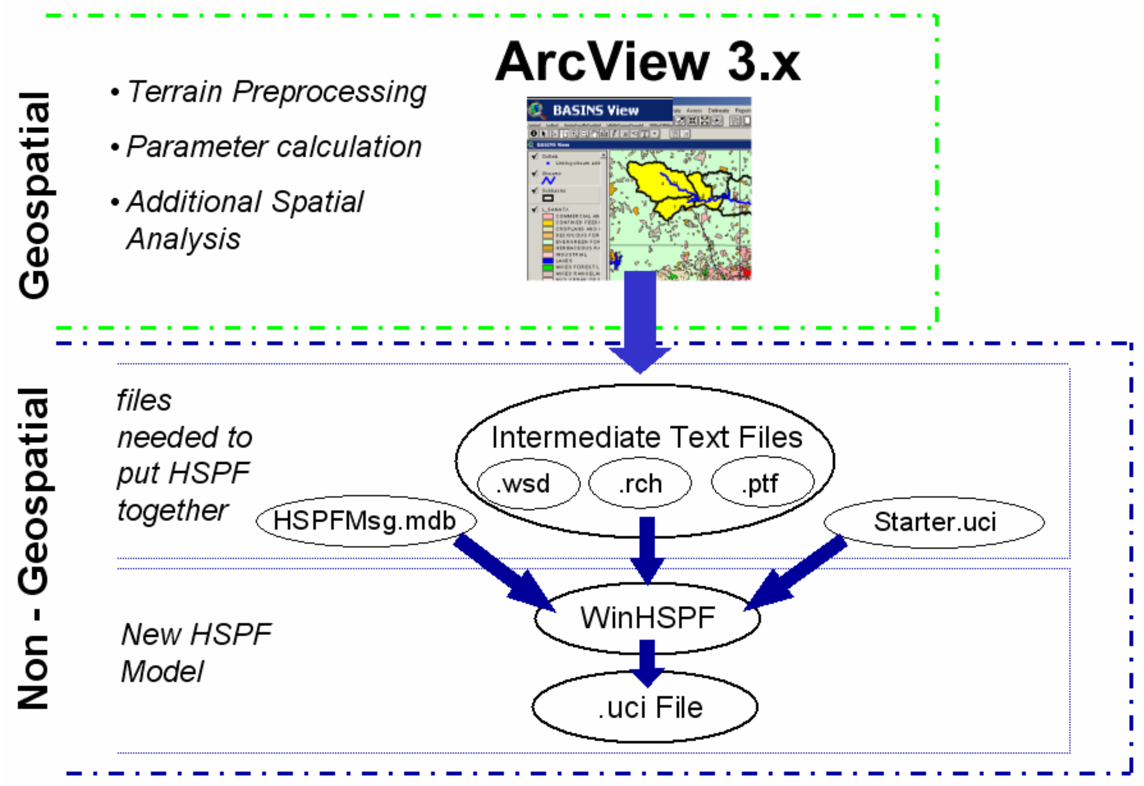


Figure 3.9 Schematic of BASINS HSPF Preprocessing methodology

Terrain Preprocessing, parameter calculation, and additional spatial analysis required to define areas to be simulated by HSPF are all performed in the GIS environment. The information necessary to define a new HSPF model is extracted to intermediate text files and combined with additional default parameters to create a new HSPF model .uci file using WinHSPF.

The BASINS HSPF Preprocessing methodology and the files it uses and produces are summarized in the following:

1) Prepare GIS data for building an HSPF model
    a. Files Used:
        i. DEM
        ii. BASINS or other external program for terrain preprocessing
    b. Tasks performed
        i. Delineate subbasins, streams, outlets
        ii. Calculate physically based parameters
    c. Files Produced
        i. Streams shapefile & attributes
        ii. Subbasins shapefile & attributes
        iii. Outlets shapefile & attributes
2) Overlay Landuse and define HSPF model configuration
    a. Files Used
        i. Streams, Subbasin, Outlets shapefiles
        ii. Landuse shapefile
        iii. hspfusgs.dbf
    b. Tasks performed
        i. Reclassify landuse categories for HSPF application
        ii. Calculate contributing area for each stream for each landuse category
        iii. Write intermediate text files
    c. Files Produced
        i. .rch, .wsd, .psr, .ptf intermediate text files
3) Create new HSPF model
    a. Files Used
        i. .rch, .wsd, .psr, .ptf intermediate text files
        ii. starter.uci text file
        iii. .wdm file with Forcing Data
        iv. (hspfmsg.mdb and hspfmsg.wdm)
    b. Tasks Performed
        i. Create .uci file from information in intermediate files
        ii. Assign Forcing Data to Operations based on initial MetSegment
        iii. Open .uci file in WinHSPF for editing
    c. Files Produced
        i. .uci file

3.5.6 HSPF MetSegments

MetSegments are a WinHSPF concept used to organize input time series for HSPF modeling. Though they are not intrinsic to the HSPF model structure they are useful for preparing input time series for a model. The MetSegment concept was developed for organizing input time series as a part of the WinHSPF interface to the HSPF model. A “MetSegment” is a conceptual model for an area of the land that receives uniform Forcing Data from the atmosphere. In the WinHSPF interface, Model Operations are associated with a MetSegment that determines which time series datasets will provide its external forcing data.

In creating an HSPF model, eight forcing time series can be assigned to a MetSegment, however, only Precipitation and Evaporation are required to run simple hydrology simulations. These eight time series are shown in Figure 3.10 and represent most of the common time series required for hydrologic and simple water quality simulations.

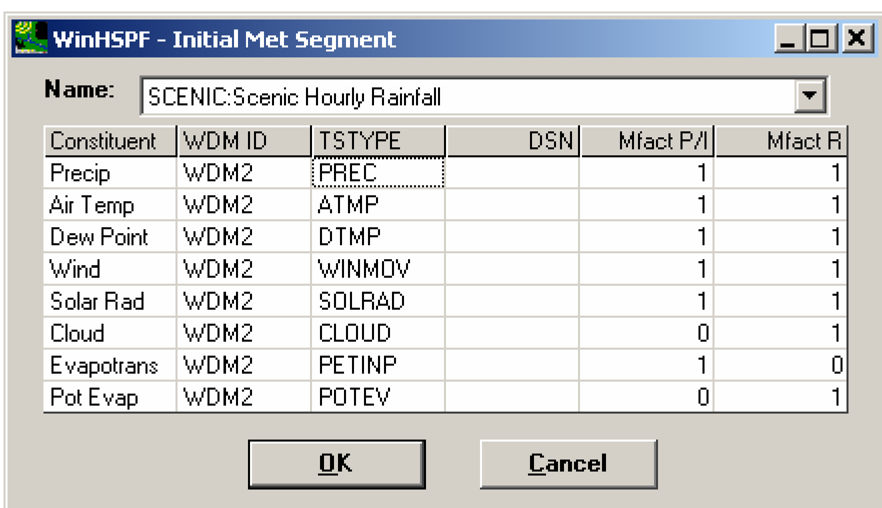


Figure 3.10 WinHSPF interface to MetSegments.

Each Operation in the HSPF model is assigned a MetSegment and, consequently, a set of forcing data corresponding to that MetSegment. In Figure 3.4, each Thiessen rain gauge area could be considered a MetSegment, and the land segments that fall within that area would be assigned the forcing time series from the appropriate MetSegment.

Some of the tools associated with MetSegments in WinHSPF are designed to have MetSegments follow Drainage Area boundaries. When a River Segment is associated with a MetSegment in WinHSPF, an option exists to assign the same MetSegment to all the land areas contributing to the River Segment. Though there is no reason to believe that areas receiving uniform rainfall actually follow drainage area boundaries, this assumption is useful if input time series are produced as "average-basin-precipitation." This concept will be discussed further in Section 3.6.

3.6 TIME SERIES PREPROCESSING FOR HSPF

Many time series are required for complex HSPF water quality models, however, this research deals primarily with setting up the configuration and preparing model files for only a simple HSPF model. The simplest HSPF models simulate only hydrologic processes (no water quality) and, if snow is not considered, only require inputs of precipitation and potential evaporation. The spatial variability of evaporation is not nearly as great as that of precipitation and HSPF contains capabilities to differentiate actual evaporation from potential evaporation based on differences in vegetative cover. Many times a single dataset describing uniform potential evaporation is assumed over the entire area simulated by even a large HSPF model. This work, therefore, is primarily concerned with developing input precipitation time series for HSPF.

3.6.1 Traditional Thiessen Polygons

Most HSPF applications use rainfall data from rain gauges located in and around the simulated area. If more than one rain gauge is present, a method similar to that presented in Figure 3.5 is typically used with Thiessen polygons or another method for assigning areas of land to the closest rain gauge. (Rainfall from the closest rain gauge is almost always used as the precipitation input to HSPF Model Operations and this idea is ultimately the motive behind the WinHSPF concept of MetSegments.) When time series from a single rain gauge is assumed to be uniform over a large area surrounding the gauge, complex spatial analysis of rainfall data is not required.

3.6.2 Interpolation Between Gauges

One alternative to simply assuming uniform rainfall over a large area is interpolating between rain gauges. Interpolation using inverse distance weighting or another algorithm can be used to estimate the rainfall for areas between gauges using the spatial analysis capabilities of a GIS. Sparse rain gauge networks, however, often do not capture the large spatial variability of rainfall on short timescales (within storm variability) and interpolation between gauges tens of kilometers apart may not be appropriate when rainfall varies on much smaller scales during a storm event.

Interpolation between gauges may give accurate results for long-term (month/annual) averages since the spatial variability of rainfall decreases over longer timescales; however, HSPF is typically run on timescales between one hour and one day. Interpolating from a sparse network of rain gauges that does not capture within-storm variability may not improve the accuracy of precipitation estimates. In addition to questions about accuracy of interpolation, complex spatial analysis of voluminous precipitation data is computationally expensive and not feasible for many applications.

3.6.3 NEXRAD Rainfall Data

An alternative to gauge rainfall data is estimates from radar, the most popular of which is NEXRAD. NEXRAD rainfall data is collected by the National Weather Service (NWS) and distributed for the entire conterminous US on a 4-km2 grid. Many data products are developed as a part of the NEXRAD program and it is beyond the scope of this paper to provide a detailed review of NEXRAD data. A data product called Stage III precipitation data represents the best estimate of rainfall available from the NWS, and from here on the term "NEXRAD data" will refer to the Stage III product unless otherwise noted. Figure 3.11 illustrates a NEXRAD dataset overlain with a watershed in GIS software.

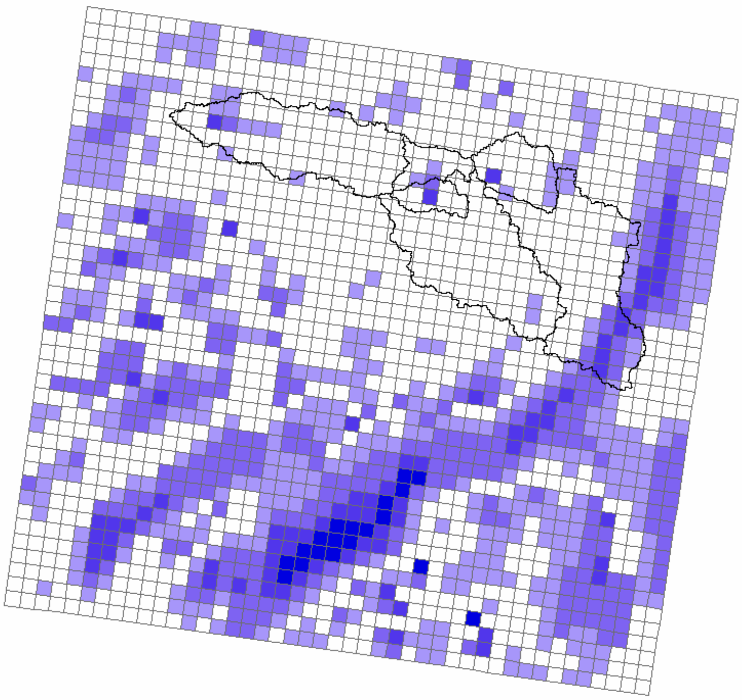


Figure 3.11 NEXRAD data, shown in ArcGIS software.

Some distributed hydrologic models, such as the Sacramento Soil Moisture Accounting (SAC-SMA) used by the NWS, are designed to make use of grid-based rainfall data such as NEXRAD (Zhang 2004). The HSPF model structure is flexible and it can be configured in many different ways to simulate complex systems. The type of rainfall data used as input to the model does influence how an HSPF model is configured but most HSPF applications are primarily structured around land surface characteristics or drainage areas. In order for NEXRAD precipitation data to be used as input to HSPF, some spatial preprocessing is necessary to make the rainfall data consistent with typical

HSPF model configurations. The most practical way for developing inputs to hydrologic models from NEXRAD data is by calculating basin-average-precipitation

3.6.4 Basin-Average-Precipitation

Many traditional "Lumped" (see footnote in Section 3.3.2) hydrologic modeling applications such as HEC-HMS or the SCS Curve Number Method rely upon estimates of basin-average-precipitation to calculate the response of an entire drainage area at a specific point on the river network. Basin here refers to the drainage area contributing to a river segment or a specific location on the river network. Basin-average-precipitation refers to an estimate of the areal average precipitation over a drainage area. Models that rely on basin-average-precipitation are described as lumped because they do not consider in detail the processes that occur within the system. They essentially "lump" these processes together and predict the response of the entire basin/river system.

Though rainfall is certainly not uniform over large drainage areas, some method must be used to estimate the average precipitation over the area. The simplest way to develop basin-average-precipitation is to assign each basin the rainfall from the closest rain gauge. More complex methods are also used to estimate basin-average-precipitation by assigning weights to surrounding gauges based on the distance from the centroid of the basin or with interpolation algorithms. Alternatively, radar estimates for rainfall can be used to estimate the basin-average-precipitation based on an area-weighted average of the grid cells that fall over the basin.

Most applications of HSPF are not "Lumped" in the traditional sense in that they do not require precipitation in a basin-average format. Hydrologic processes are typically simulated using one or only a few precipitation datasets on several land surface types. The results of these simulations are distributed (based on contributing area) to the various segments of the river network to estimate its response. Though the HSPF model structure does not require basin-average-precipitation, a model could be configured to make use of rainfall data in a basin-average format. Figures 3.12 and 3.13 show an HSPF model configuration that is capable of receiving rainfall time series developed as basin-average-precipitation.

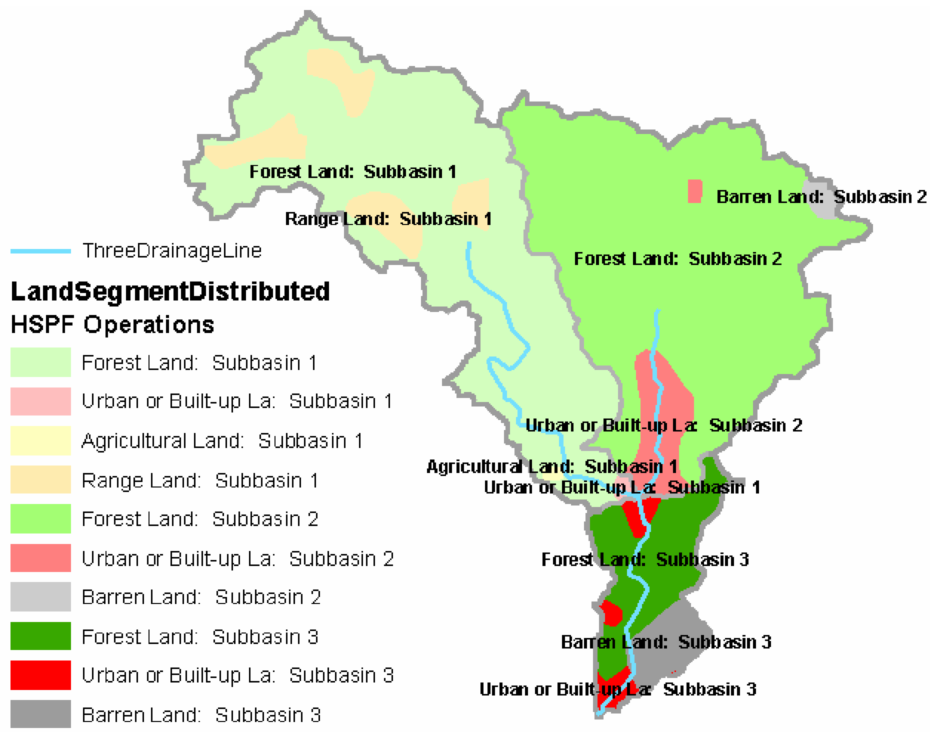


Figure 3.12 HSPF configuration for basin-average-precipitation.

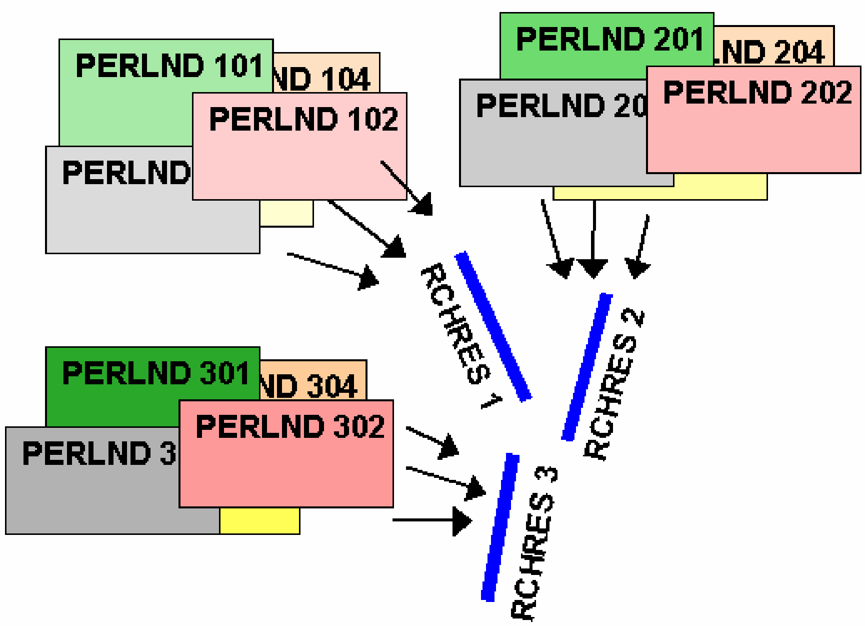


Figure 3.13 Schematic of HSPF basin-average-precipitation model configuration.

In this configuration, a single set of Land Segments (divided according to landuse categories) is used to simulate the drainage area associated with each Reach Segment in the model. This configuration could be used to assign precipitation time series developed in a basin-average way to each River Segment and its associated Land Segments.

Many HSPF applications are not configured to make use of rainfall time series developed as basin-average-precipitation because of the computational expense of simulating hundreds of Land Segments. If a unique set of Land Segments is created for each Reach Segment drainage area in a large HSPF model, hundreds of Land Segments must be simulated. If, however, uniform rainfall can be assumed, the entire land area could be simulated with only a few Land Segments. Fewer Model Segments mean less time to perform a simulation, and less complexity in the .uci files.

Another reason that HSPF models use "Lumped" representations (Figure 3.3 vs. Figure 3.13) of the land surface is because distributed rainfall data is not widely available in a convenient format. Time series from gauges is easy to obtain from the National Weather Service and other sources. It is often distributed in a format that nearly matches the .wdm file format and can be easily used as inputs to the HSPF model. This time series can be directly input to an HSPF model, or simple adjustments can be used to make estimates for areas between gauges. NEXRAD data, on the other hand, is distributed in a gridded format, far from that required as input to HSPF models. An individual file is often used to represent the rainfall for each timestep, meaning the data need to undergo a space-time recompositioning process in order to be compatible with the .wdm file format used for HSPF modeling. This data manipulation involves extracting values from potentially thousands of gridded rainfall files and writing them to a traditional timeseries format (See Section 2.5). Because HSPF is not designed to make use of distributed rainfall data, some computationally expensive spatial analysis of gridded datasets is also necessary to make precipitation data compatible with the chosen HSPF model configuration.

3.6.5 Grouped vs. Individual Option in WinHSPF

Two options are presented at the end of the BASINS HSPF Preprocessing tools in Section 3.5.1, one for creating 'Grouped' Land Segments, and another for creating 'Individual' Land Segments. This option is used to inform WinHSPF what type of model configuration to use when building the new .uci file. The "Grouped" option creates a single set of Land Segments (divided according to lanudse categories) to simulate the entire area of interest. Each Land Segment Operation contributes to the appropriate River Segment in proportion to the amount that falls in each Drainage Area. The 'Grouped' option is used if a single time series of precipitation will be used as input to all the area simulated by the model. Section 3.3 (and specifically Figure 3.3) discusses the implications and limitations of this type of model configuration.

The "Individual" option results in a model configuration in which a single set of Land Segments (divided according to landuse categories) is created for the drainage area associated with each Reach Segment in the model. Precipitation time series are assigned to Model Elements in the .uci file by associating each River Segment with an appropriate MetSegment (timeseries datasets) in WinHSPF and using tools to apply the same MetSegment to all to the Land Segments contributing to the River Segment. This configuration is presented in Figure 3.13 and inherently assumes that all the land in a River Segment's drainage area will receive uniform precipitation, or "basin-average-precipitation."

Though the BASINS WinHSPF tools are designed to work with time series in a basin-average way, rainfall time series are typically not developed in this manner because of the difficulties associated with preparing data for input to the model. The time and resources required to develop true estimates of basin-average precipitation for input to a configuration such as that shown in Figure 3.13 are prohibitive for many HSPF applications. Because of these limitations, point measurements of rainfall from nearby gauges are typically assumed to be uniform over large areas simulated by HSPF models.

3.7 SUMMARY

Geospatial information is used in HSPF modeling primarily during model development. GIS tools and data are useful for calculating physically based attributes and for defining areas of land with similar hydrologic characteristics. The HSPF model structure is general enough to allow flexibility in the configuration chosen to simulate a system of interest. Though HSPF models are typically structured around differences in land surface characteristics identified using GIS data, the type of forcing data used in the simulation also affects the configuration used to simulate a system.

Chapter 4 Methodology

Chapters 2 and 3 provide a summary of the state of knowledge and current approaches for developing HSPF models. A thorough understanding of the working of the HSPF model, its data types, and different possibilities for configurations, are all important for implementing a successful methodology for preprocessing GIS data for the model. In addition, existing tools from the BASINS and Arc Hydro systems provide a foundation for new HSPF model development tools.

The purpose of this chapter is to present an overview of a methodology designed to create new HSPF models in the ArcGIS environment. The work presented here is organized into two sections. The first is an ArcGIS HSPF Preprocessing methodology designed to facilitate HSPF model development in ESRI’s Arc9 ArcToolbox environment. It essentially mirrors that of the BASINS HPSF Preprocessing methodology presented in Chapter 3, but implements it in the ArcGIS environment. The second section of this chapter presents a methodology for using ArcGIS data to develop input time series for HSPF. The ArcGIS Timeseries Preprocessing methodology uses the Arc Hydro time series structure and the BASINS/WinHSPF concept of MetSegments to prepare time series and model files.

Table 4.1 presents an overview of the tools used to implement the methodologies presented in the following sections. Major tasks and their integration are outlined, but detailed explanations of the specific data structures and tools used to implement the methodologies are provided in the next chapter.

Table 4.1 Overview of ArcGIS HSPF and Timeseries Preprocessing methodologies

	Tasks required to create new HSPF model	Arc Hydro Data and Tools	ArcGIS Tools	WinHSPF/GenScn/ WdmUtil
ArcGIS HSPF Preprocessing methodology	Define drainage area boundaries and river network	Arc Hydro Catchment/ DrainageLine data		
ArcGIS HSPF Preprocessing methodology	Calculate Physically-based parameters		Standard ArcGIS tools*	
ArcGIS HSPF Preprocessing methodology	Define Model Elements to be simulated		Standard ArcGIS tools*	
ArcGIS HSPF Preprocessing methodology	Extract data from GIS to intermediate text files		Custom ArcGIS tools**	
ArcGIS HSPF Preprocessing methodology	Create new .uci file from intermediate text files			WinHSPF 'Create Project' tool
ArcGIS Timeseries Preprocessing methodology	Write timeseries data to .wdm file	TimeSeries structure	Custom ArcGIS tools**	WdmUtil/GenScn programming libraries
ArcGIS Timeseries Preprocessing methodology	Update .uci file to read from new timeseries datasets		Custom ArcGIS tools**	WdmUtil/GenScn programming libraries

**Though standard ArcGIS tools are available for these tasks, ModelBuilder models were created to streamline the process.*

***Custom ArcGIS Geoprocessing functions were created to perform these tasks.*

4.1 ArcGIS HSPF Preprocessing Methodology

As shown in Table 4.1, the ArcGIS HSPF Preprocessing methodology requires the use of tools provided by Arc Hydro, ArcGIS, and WinHSPF. In addition, a number of custom tools were developed as a part of this research. The Arc Hydro data model and terrain processing tools are widely used in the hydrologic modeling community for defining and organizing basic data in the ArcGIS environment. The Catchment and DrainageLine feature classes are the result of the Arc Hydro terrain processing tools and contain the spatial information necessary to define the drainage areas of each segment of the river system and the network information necessary to define the connectivity in the watershed and river network system. The ArcGIS HSPF Preprocessing methodology is designed to utilize data for river networks and associated drainage areas resulting from the application of the Arc Hydro terrain preprocessing tools.

It is important to note that weather or not the Arc Hydro terrain processing tools are used to develop stream and drainage area data, the Arc Hydro data model provides a general, robust means of organizing data from any source while maintaining a few, essential relationships. Thus, the use of the well-developed data structure provided by Arc Hydro avoids the necessity of additional ArcGIS terrain processing tools.

In addition to making use of the existing Arc Hydro tools and data model, the ArcGIS HSPF Preprocessing methodology also makes use of WinHSPF, a non-GIS component of the BASINS system. In the initial development of the BASINS system, an effort was made to separate tasks that do not require complex spatial analysis from those that require GIS capabilities. One result of this effort is the WinHSPF software package which operates independently of GIS software and uses only essential information extracted from GIS data to define the structure of a new HSPF model.

WinHSPF is commonly used in initial HSPF model development following spatial analysis with the ArcView 3.x components of the BASINS system. However, because its algorithms require no complex spatial analysis, WinHSPF can be used to build a new .uci file independently of the GIS components of the BASINS system provided the required data is available from another source. The ArcGIS HSPF Preprocessing methodology makes use of the capabilities of WinHSPF by extracting information from ArcGIS data and converting it to the format required to build a new HSPF model.

The ArcGIS methodology developed in this research essentially mirrors the BASINS HSPF Preprocessing methodology, and the only major conceptual difference concerns the attempt to maintain a geospatial description of model elements in the GIS data. This concept is described in the Section 4.3 and will be used after initial model development to facilitate the transfer of information between GIS data and HSPF model files.

Figure 4.1 illustrates the components of the ArcGIS and BASINS HSPF Preprocessing methodologies. While the BASINS and ArcGIS HSPF Preprocessing methodologies use different GIS tools, they both use WinHSPF to build a new HSPF model using information extracted from GIS data.

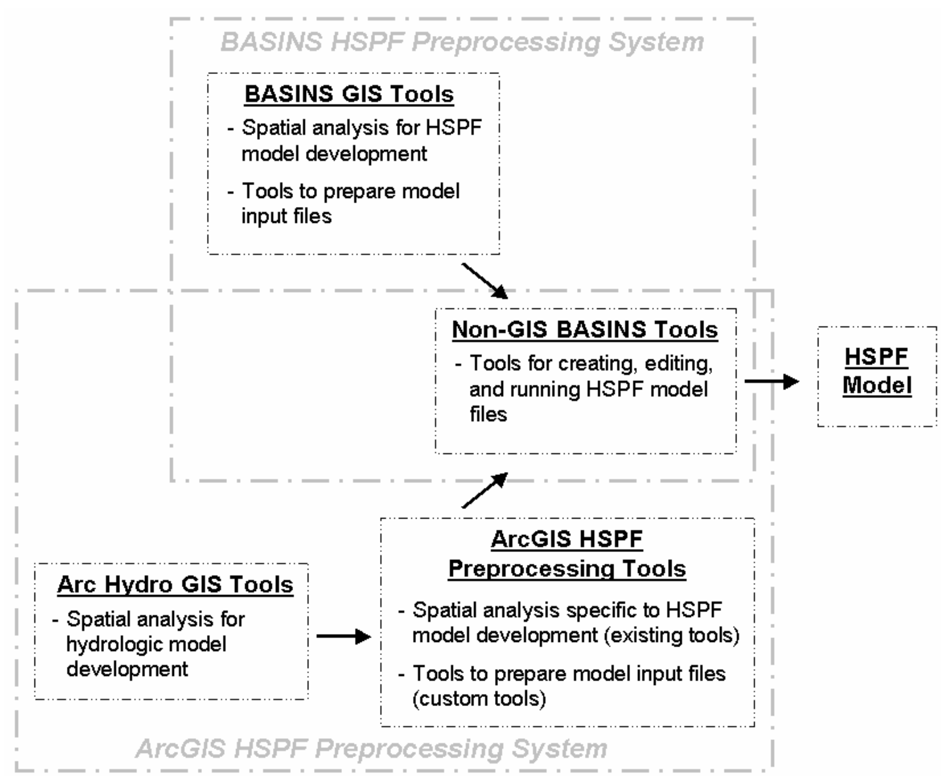


Figure 4.1 Schematic overview of ArcGIS and BASINS HSPF Preprocessing methodologies.

The following sections introduce the major tasks accomplished by the ArcGIS HSPF Preprocessing methodology in developing a new HSFP model.

1) Utilize Arc Hydro (or equivalent) data for drainage areas boundaries and river networks
2) Calculate physically-based attributes required for HSPF model creation
3) Define River and Land Segments (in GIS) to be simulated by an HSPF model
4) Write intermediate text files to utilize the capabilities of WinHSPF to create a new HSPF model
5) Create a new .uci file using the WinHSPF program

4.1.1 Drainage Area Boundaries and Stream Networks: Arc Hydro Data

The results of the basic "Terrain Processing" tools in Arc Hydro are feature classes describing Drainage Lines, Catchments, and Drainage Points. Figure 4.2 shows the result of the application of the Arc Hydro Terrain Processing tools. Though the ArcGIS HSPF Preprocessing methodology is designed to use data for from the Arc Hydro data model, data from another source can be used as well, provided some essential information is present.

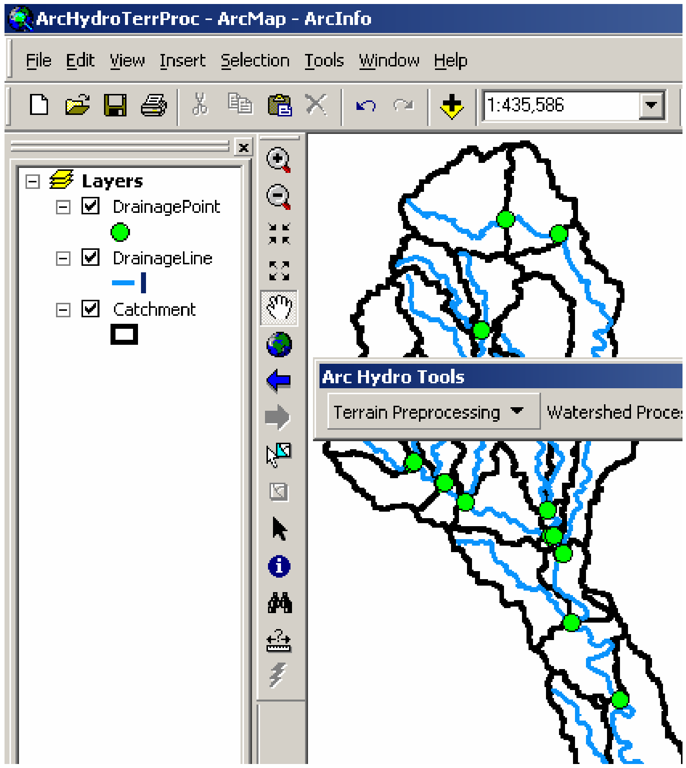


Figure 4.2 GIS results of Arc Hydro terrain preprocessing tools.

The three Feature Classes shown in Figure 4.2 contain attributes summarized in Table 4.2. Though the attribute names shown in Table 4.2 are specific to the Arc Hydro Data Model, the information they contain is general and not specific to any type of application or terrain processing methodology.

Table 4.2 Attributes from Arc Hydro terrain preprocessing tools

FeatureClass	***Attribute***	***Description***
DrainageLine	HydroID	Unique feature identifier in the geodatabase
DrainageLine	HydroCode	Permanent public identifier of the featuer
DrainageLine	DrainID	HydroID of the reference drainage area feature
DrainageLine	Shape_Length	
Catchment	HydroID	Unique feature identifier in the geodatabase
Catchment	HydroCode	Permanent public identifier of the featuer
Catchment	DrainID	HydroID of the reference drainage area feature
Catchment	AreaSqKm	Area in square kilometers
Catchment	JunctionID	HydroID of the HydroJunction at drainage outlet
Catchment	NextDownID	HydroID of the next downstream catchment
Catchment	Shape_Area	
DrainagePoint	HydroID	Unique feature identifier in the geodatabase
DrainagePoint	HydroCode	Permanent public identifier of the featuer
DrainagePoint	DrainID	HydroID of the reference drainage area feature
DrainagePoint	JunctionID	HydroID of the HydroJunction at drainage outlet

The attributes of the Arc Hydro data presented in Table 4.2 contain the network information necessary to begin defining the configuration of an HSPF model. Additionally, the polygon Catchment boundaries define the drainage areas necessary to define the contributing areas of Land Segments in HSPF.

4.1.2 Calculate Physically-Based Parameters

The Arc Hydro tools possess the spatial analysis capabilities to delineate river networks and drainage areas but additional analysis is required to calculate physically-based parameters for HSPF modeling. The ArcGIS HSPF Preprocessing system uses existing ArcGIS tools to calculate the physically-based parameters for HSFP modeling and stores them as attributes of feature classes.

Of the attributes necessary to use the BASINS HSPF Preprocessing system to build a new .uci file (shown in Table 3.2), only seven of them are physically based and require some sort of spatial analysis. The ArcGIS HSPF Preprocessing system uses only seven physically-based attributes and they are presented in Table 4.3.

Table 4.3 Physically-based attributes required for HSPF model development

Description	***Associated GIS Data***	***How Calculated***
Stream Length	Streams	GIS Shape Length
Min Elevation of Stream	Streams	Stream Feature and DEM
Max Elevation of Streams	Streams	Stream Feature and DEM
Stream Slope (average)	Streams	Calculated from Length, Max, Min Elevations
Steram Width (average)	Streams	Extrenal Program or Information
Stream Depth (average)	Streams	Extrenal Program or Information
Subbasin Slope (average)	Subbasins	Subbasin polygons and DEM/Slope Raster

For streams, maximum and minimum elevations are combined with length to calculate the average slope for the stream. The methodology assumes that one (and only one) feature is present in GIS data for each stream to be simulated with HSPF. Figure 4.3 illustrates the process of calculating stream slope.

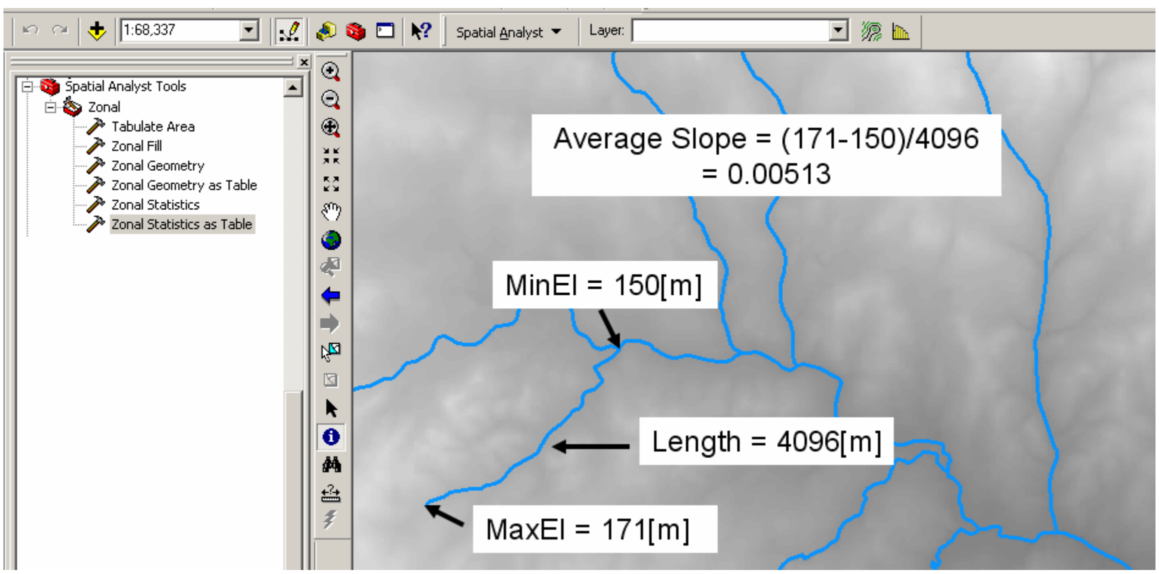


Figure 4.3 Calculating average stream slope.

This method for calculating the stream slope is appropriate for most cases when the longitudinal slope of a stream does not change significantly within the reach. If the longitudinal slope changes dramatically, more accurate model results can be obtained by dividing the reach into two or more pieces.

Average stream width and depth are also used by the WinHSPF ‘Create Project’ tool to produce a description of channel geometry in the .uci file. The details of an HSPF model configuration are not dependent upon an accurate description of river geometry, and the ArcGIS HSPF Preprocessing methodology does not contain any tools for estimating stream width and depth. If values are not available from another source, a default width and depth is assumed, and a more accurate description of river geometry can be added at a later time.

For the drainage areas, slope is the only physically based attribute for which GIS calculations are required. Average slope for a drainage area can be calculated from a GIS raster dataset using ArcGIS Zonal Statistics on a polygon feature class describing the Subbasin areas. If a slope dataset is not available, ArcGIS’s ArcToolbox contains tools that can be used to create a slope raster from DEM data.

Those familiar with BASINS HSPF Preprocessing will notice the absence of one additional physically-based attribute Overland flow distance (LSUR). Though it is stored on the BASINS Subbasins shapefile as an attribute “Len1,” it is not ultimately used by WinHSPF in HSPF model creation and was omitted from the ArcGIS system.

4.1.3 Define HSPF Model Elements

After calculating physically-based parameters using ArcGIS tools, Model Elements to be simulated by HSPF must be defined. River Segments are simulated as a lumped, completely mixed system by the HSPF model. Drainage lines resulting from the Arc Hydro Terrain Processing tools have a one-to-one relationship to the catchments containing them, and each feature in the Arc Hydro DrainageLine feature class essentially is an explicit representation of HSPF RCHRES Model Elements. Network information, stream slope, width, depth, and length stored as attributes of GIS data can be used directly to define HSPF RCHRES Model Elements. Defining Land Segments, however, is more complicated.

The first step in defining Land Segments for HSPF modeling is choosing the types of Landuse that will be simulated in the HSPF model. These types must then be mapped to Landuse categories defined in the GIS data. This task is accomplished behind the scenes in the BASINS HSPF Preprocessing system using the 'hspfusgs.dbf' table. In the ArcGIS HSPF Preprocessing methodology, this task is similarly accomplished with a table stored in a geodatabase.

Figure 4.4 shows polygon Landuse data (Landuse feature class) with a unique ID representing the Level 2 Anderson Classification categories and a table (LuseDefinition table) used to map these categories to types to be simulated by HSPF.

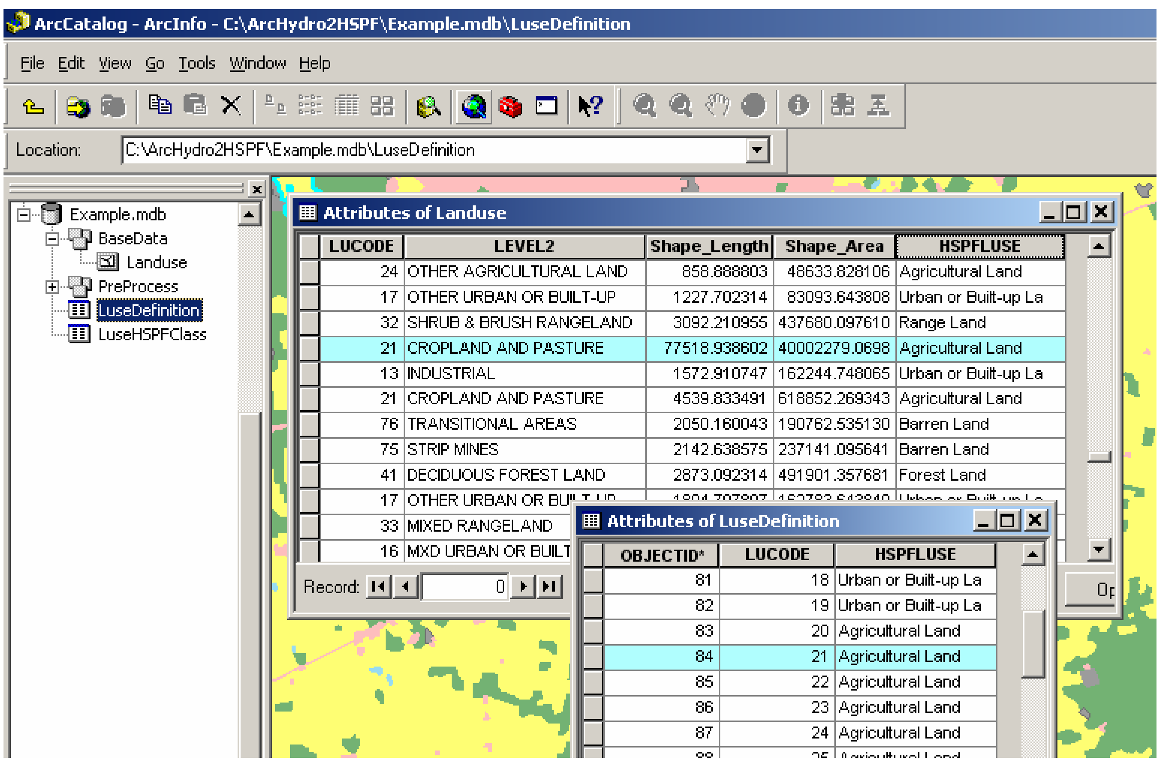


Figure 4.4 Defining landuse types to be simulated by HSPF.

Once the types of land to be simulated by HSPF have been defined in GIS data, additional spatial analysis is necessary to calculate the amount of each type of land that contributes to each River Segment. Standard ArcGIS topological functions are available to make these calculations using the Arc Hydro drainage area boundaries and landuse data. Figures 4.5 and 4.6 illustrate the results of the application of the ArcGIS tools.

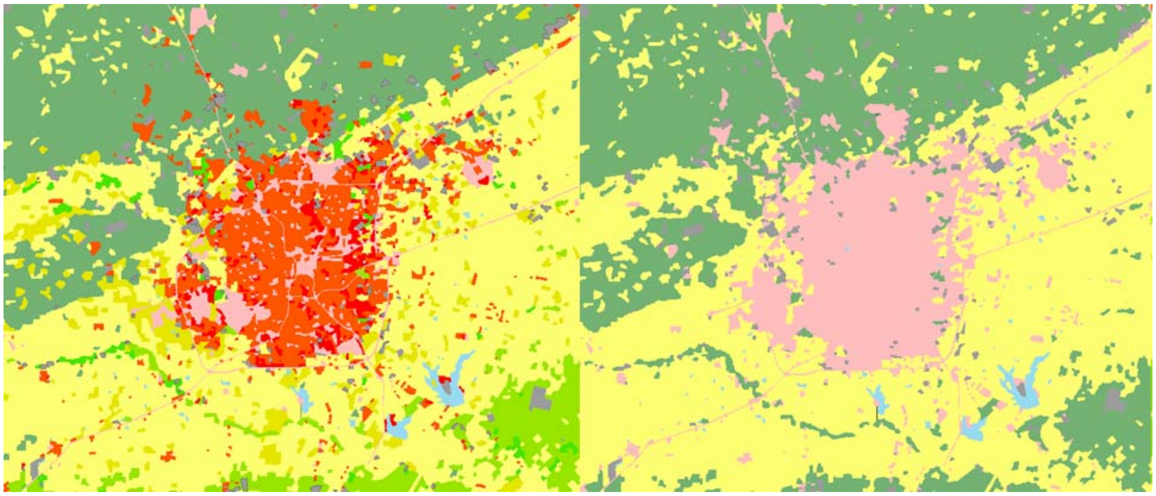


Figure 4.5 ArcGIS dissolve function – Spatial view.

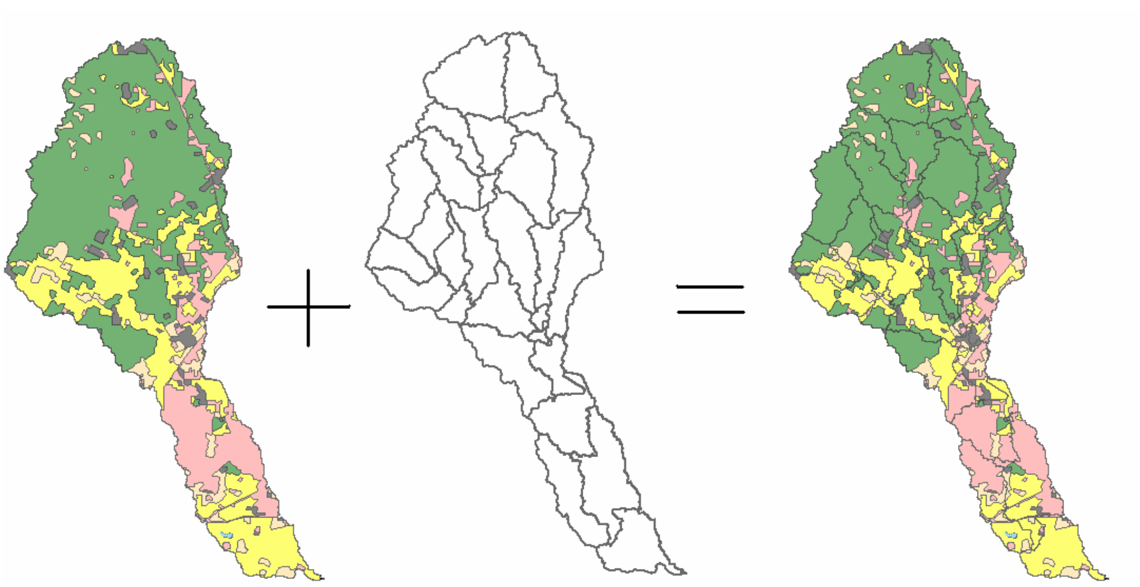


Figure 4.6 ArcGIS intersect function – Spatial view.

Figure 4.5 illustrates the process of defining the types of landuse to be simulated by HSPF. GIS data commonly contains 30+ categories of land (figure at left), but HSPF applications typically simulate fewer than 10 types of land (figure at right). ArcGIS tools can be used to map GIS landuse categories to HSPF landuse types.

Figure 4.6 illustrates the process of calculating the amount of each landuse type that contributes to each river segment. Drainage area boundaries are overlain with landuse data to find the landuse distribution in each. The result of the GIS function illustrated in Figure 4.6 is a polygon feature class in which each feature represents the area of a single landuse type that contributes to a single river segment.

The process outlined above makes use of polygon landuse data, however, raster landuse data are also widely used in HSFP model development. The procedure of Application chapter presents the ArcGIS functions used to calculate contributing areas for polygon landuse data, and Appendix A presents a methodology for using raster landuse data.

Standard ArcGIS functions are used to calculate the distribution of landuse types for each drainage area; however, further consideration must be made in defining Impervious Land Segments. In the BASINS HSPF Preprocessing methodology, Impervious Land Segments are defined by manually assigning a "Percent Pervious" to each Landuse category to be simulated in HSPF. For each category that contains a "Percent Pervious" less than 100%, two contributing areas are defined for each River Segment, one for Pervious Land and one for Impervious Land.

In the ArcGIS HSPF Preprocessing system, a similar procedure is used and the process is automated by storing the "Percent Pervious" in a table in the geodatabase. The table shown in Figure 4.7 assigns a "Percent Impervious" to each type of landuse to be simulated by HSPF rather than a Percent Pervious.

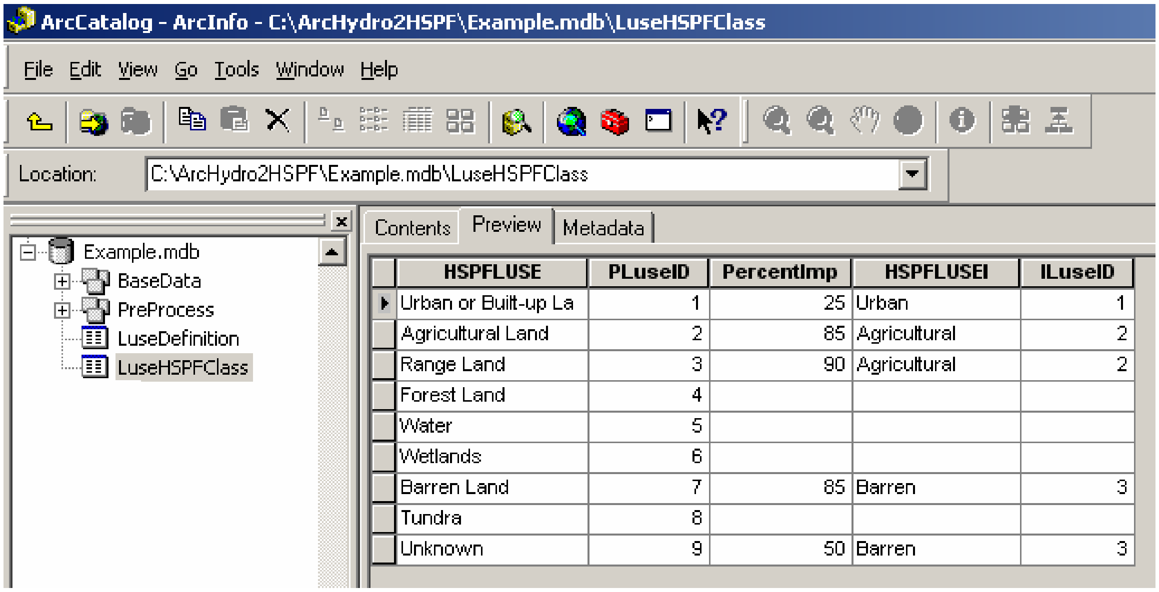


Figure 4.7 ArcGIS table to define impervious area, operation numbering convention.

The only difference between the algorithm implemented in the ArcGIS HSPF Preprocessing tools and the BASINS process is that pervious landuse types can be combined when defining impervious types. Figure 4.7 shows that the impervious land from both Range Land and Agricultural Land are grouped into the same Impervious “Agricultural Land” category. This algorithm is slightly more general than the one implemented in the BASINS HSPF Preprocessing system since it allows an impervious landuse type to include areas from one or more pervious landuse type.

4.1.4 Extract Data to Intermediate Text Files

The ArcGIS functions presented above prepare GIS data in such a way that information necessary for HSPF model development can be easily extracted and written to intermediate text files. Though standard ArcGIS tools are used to perform the previous functions, custom ArcGIS tools have been developed to implement the process of extracting data to the intermediate text files. These tools are described in detail in the

Chapter 5, and a table outlining the method used to create the text files is presented in Appendix D.

4.1.5 Create New .uci File

After necessary information has been extracted from GIS data and written to intermediate text files with custom ArcGIS tools, the WinHSPF program is used to create a new .uci file outside the GIS environment. This tool is accessed through the non-GIS WinHSPF software and uses information from intermediate text files produced by either the BASINS or ArcGIS Preprocessing tools.

4.2 ARCGIS TIMESERIES PREPROCESSING METHODOLOGY

The second part of this research involves the use of ArcGIS data and tools after initial HSPF model development. The ArcGIS Timeseries Preprocessing methodology uses ArcGIS to organize and prepare timeseries data and updates HSPF model files to read timeseries from the appropriate data sets.

Though no spatial information is stored explicitly in the .uci file, each Model Element simulated by HSPF represents some spatial location in the real world. In order to facilitate the transfer of information from GIS data and HSPF model files after initial model development, a spatial representation of the areas of land simulated by HSPF must be maintained. In order to implement the ArcGIS HSPF Preprocessing methodology, GIS data used in model development are stored in a geodatabase and used to locate Model Elements after initial model development.

The methodology was designed to make use of NEXRAD precipitation data which is commonly stored in a format very different from the .wdm file format required for HSPF modeling. Appendix B describes the process used to extract NEXRAD data from its native binary format and write it to the Arc Hydro timeseries format. The

ArcGIS Timeseries Preprocessing methodology presented here assumes that NEXRAD data is already stored in a geodatabase following the Arc Hydro timeseries structure.

4.2.1 ArcGIS Timeseries Preprocessing Overview

The ArcGIS Timeseries Preprocessing methodology is fundamentally built upon two concepts:

1) the Arc Hydro relational database time series structure (See section 2.5)
2) an explicit geospatial representation of HSPF Model Elements in GIS data (explained further in Section 4.3)

Arc Hydro timeseries values are attached to spatial features, and consequently, GIS data can be used to locate precipitation data in the real world if it stored using Arc Hydro. The ArcGIS Timeseries Preprocessing methodology attempts to maintain a geospatial representation of HSPF Model Elements in GIS data, so that each river segment and area of land simulated by HSPF can be located in the real world. With a spatial representation of the locations of precipitation data and HSPF Model Elements, GIS tools are used to transfer information from GIS data to model files

A GIS representation of the WinHSPF concept of MetSegments is used to organize and prepare input time series in the ArcGIS Timeseries Preprocessing system. This GIS MetSegment feature class is the central data structure used to implement the ArcGIS HSPF Preprocessing methodology. A more detailed description of the MetSegment concept can be found in Section 3.4.6 and the WinHSFP user's manual. (Duda 2001)

Figure 4.8 illustrates the overall concept of the ArcGIS Timeseries Preprocessing methodology. At the center of the ArcGIS Timeseries Preprocessing system is the GIS MetSegment feature class. It contains information to communicate with both Arc Hydro timeseries data and HSPF model files. GIS MetSegments contain attributes to communicate directly with .wdm files, but the link to the .uci file is made through the geospatial representation of HSPF Model Elements.

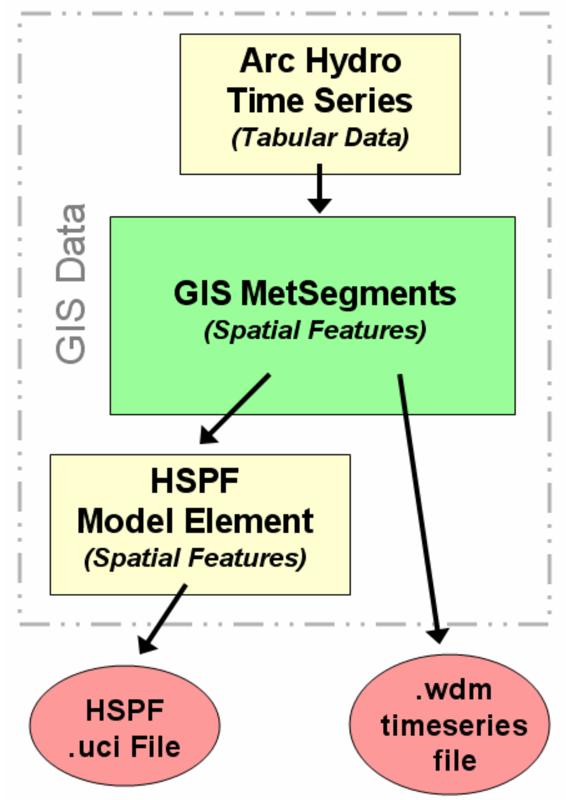


Figure 4.8 Schematic overview of ArcGIS Timeseries Preprocessing methodology.

The ArcGIS Timeseries Preprocessing methodology accomplishes two tasks to prepare precipitation data for HSPF modeling:

1) use time series data stored in a geodatabase to write time series to .wdm data sets

2) update the model .uci file to read this time series from the appropriate data sets

4.2.2 Writing Time Series to .wdm File

Each feature of the GIS MetSegment feature class contains attributes specifying data set numbers to write forcing data for HSPF modeling. The ArcGIS Timeseries

Preprocessing methodology uses timeseries values attached to these GIS MetSegments, in the Arc Hydro timeseries format, to write .wdm file data sets.

A custom ArcGIS Geoprocessing tool loops through each feature in the GIS MetSegment feature class and extracts the appropriate timeseries values. These values are then written to .wdm data set location specified in the attributes of the GIS MetSegment feature class. The details of this process are described in Chapter 5.

4.2.3 Updating HSPF Model .uci File

The second component of the ArcGIS Timeseries Preprocessing system involves updating the model .uci file so that Model Elements receive precipitation and other forcing data from the appropriate .wdm dataset. The tools developed to accomplish this task use a geospatial representation of Model Elements to locate the Model Element in the GIS data.

A custom ArcGIS Geoprocessing tool loops through each Model Operation (PERLND, IMPLND, and RCHRES) in the .uci file and finds the corresponding GIS feature in the geodatabase. A relationship between the GIS representation of the Model Element and GIS MetSegments is then used to assign the appropriate time series datasets in the .uci file as specified in the geodatabase. Maintaining a geospatial representation of HSPF Model Elements in GIS data and facilitating the transfer of information between .uci file and GIS data is not an easy task. The method used to implement the ArcGIS Timeseries Preprocessing methodology is the subject of Section 4.3.

4.3 MAINTAINING A GEOSPATIAL REPRESENTATION OF HSPF MODEL ELEMENTS

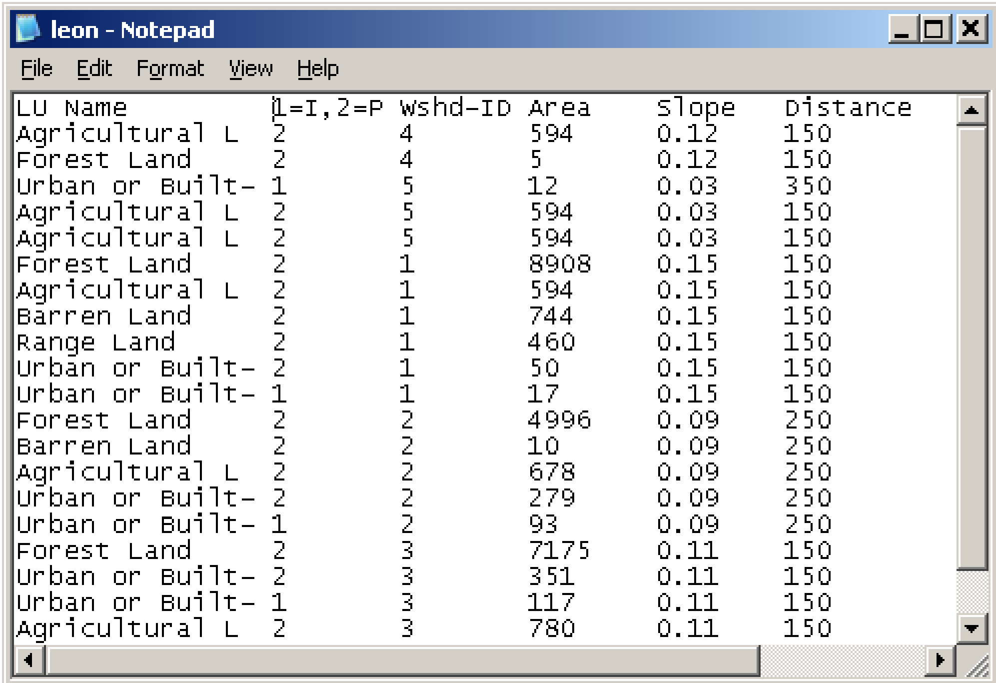
In order to facilitate the transfer of information from GIS data to HSPF model files after initial model development, a geospatial representation of Model Elements must be maintained in GIS data. In both the BASINS ArcView 3.x and the ArcGIS

Preprocessing system, information is extracted from GIS data and written to intermediate text files during initial model development. At this point, all explicit geospatial information about the areas of land simulated by HSPF is lost and Land Segments are only represented by numbers storing contributing areas and schematic information. In the BASINS HSPF Preprocessing methodology, no attempt is made to save a geospatial representation of the lines of the .wsd text file.

In the ArcGIS HSPF Preprocessing methodology, an attempt is made to maintain GIS data representing the land areas simulated by HSPF. Land segments in HSPF models can be configured in many different ways and, to maintain consistency, in this research a specific, well-defined model configuration has been chosen. In order for the geospatial representation of HSPF Model Elements to be explicit, the configuration defined by WinHSPF when the "Individual" option is used to create a new .uci file is adopted as a standard.

4.3.1 .wsd File Lines and the "Individual" and "Grouped" Options in WinHSPF

The lines of the .wsd intermediate text file contain information from GIS data used to define the Land Segments to be simulated by HSPF. Figure 4.9 illustrates the lines of a .wsd file. Each line of a .wsd file defines the area of a single type of land that contributes to a single HSPF River Segment. Separate lines are present for pervious and impervious land segments.


leon - Notepad

File Edit Format View Help

LU Name	1=I,2=P	Wshd-ID	Area	Slope	Distance
Agricultural L	2	4	594	0.12	150
Forest Land	2	4	5	0.12	150
Urban or Built-	1	5	12	0.03	350
Agricultural L	2	5	594	0.03	150
Agricultural L	2	5	594	0.03	150
Forest Land	2	1	8908	0.15	150
Agricultural L	2	1	594	0.15	150
Barren Land	2	1	744	0.15	150
Range Land	2	1	460	0.15	150
Urban or Built-	2	1	50	0.15	150
Urban or Built-	1	1	17	0.15	150
Forest Land	2	2	4996	0.09	250
Barren Land	2	2	10	0.09	250
Agricultural L	2	2	678	0.09	250
Urban or Built-	2	2	279	0.09	250
Urban or Built-	1	2	93	0.09	250
Forest Land	2	3	7175	0.11	150
Urban or Built-	2	3	351	0.11	150
Urban or Built-	1	3	117	0.11	150
Agricultural L	2	3	780	0.11	150

Figure 4.9 .wsd intermediate text file structure.

The way in which the lines from the .wsd file are used to create Land Segments in the new .uci file defines the HSPF model configuration. The WinHSPF “Individual” option for creating a new .uci file creates a single HSPF Land Segment for each line of the .wsd file, resulting in a unique set of land segments (one for each type) for each drainage area. An example of this configuration is shown in Figure 3.13. The “Grouped” option creates only a single HSPF Land Segment for each type of landuse encountered in the .wsd file, resulting in a single set of land segments (one for each type) that contribute to the appropriate River Segment based on drainage area. An example of this configuration is presented in Figure 3.3.

If impervious land is ignored for the moment and the “Individual” option for creating a new .uci file with WinHSPF is assumed, the final landuse feature class presented in Section 4.1.3 (Figure 4.6) contains features representing the lines of the .wsd file and, consequently, a representation of the (pervious) Land Segment Model Elements to be simulated by HSPF. Figure 4.10 shows this feature class, in which each feature contains information describing its landuse type, its subbasin (equivalently, which river segment it contributes to), and its surface slope.

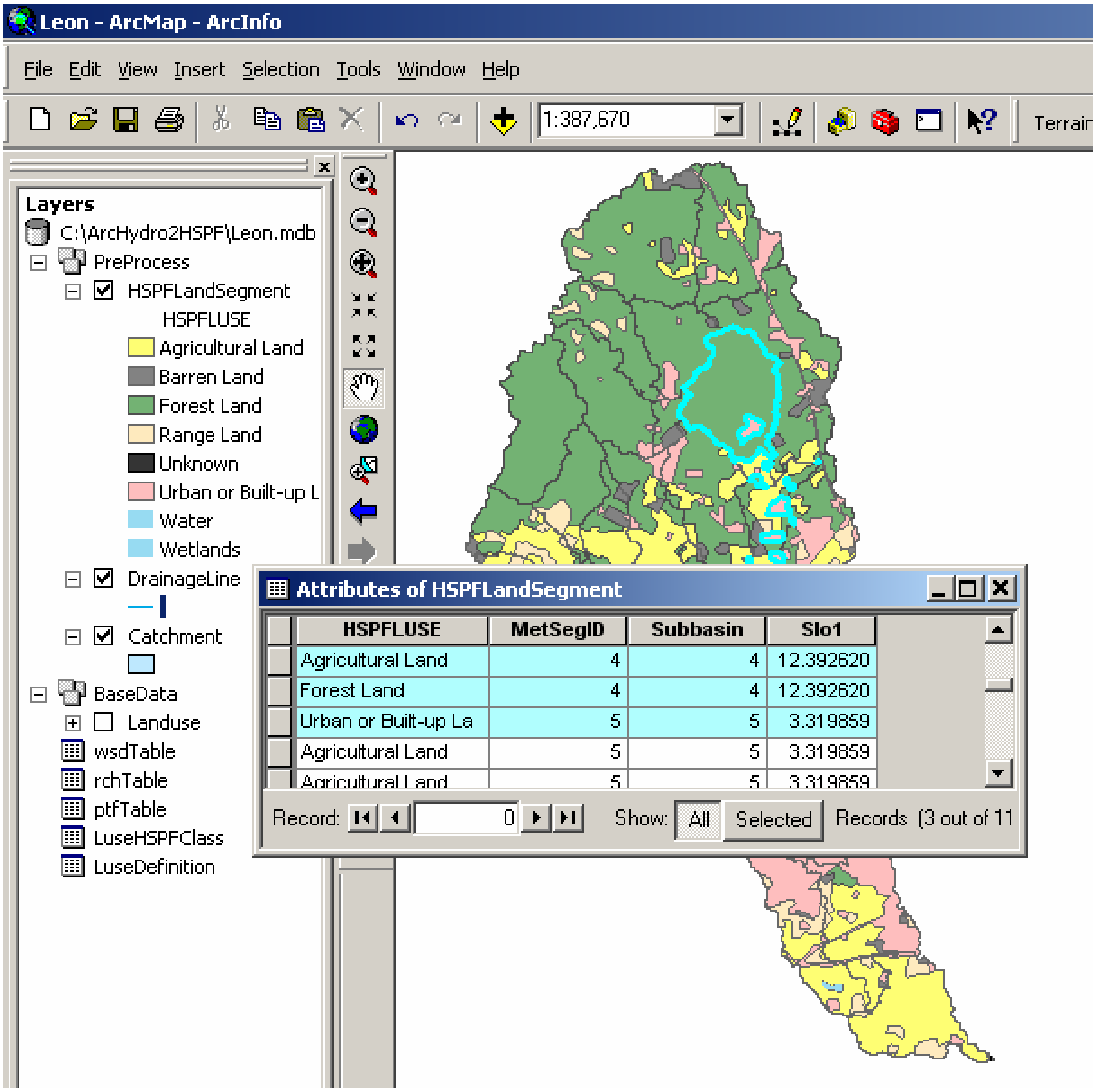


Figure 4.10 GIS data used to defining HSPF Model Elements.

Only the tabular attributes of these features are used to write the .wsd file and the actual geospatial information (description of the polygon areas) is not necessary for model creation. However, maintaining a geospatial representation of Model Elements allows an HSPF modeler the option of using GIS data to explicitly locate the Model Element in the real world.

With the ability to locate HSPF Model Elements in space using GIS data, other GIS tools and data can be used after initial model development to automatically assign timeseries data sets – as in the ArcGIS HSPF Preprocessing methodology – or for the visualization of results, model calibration, data management, or to transfer data between different models using GIS as a common platform. Additional applications of GIS data to HSPF modeling are discussed in Chapter 7.

4.3.2 Complications Associated with Maintaining Geospatial Information

If impervious landuse is not considered, and the WinHSPF "Individual" option is used to create a new .uci file, the feature class described above contains an explicit representation of each Land Segment simulated by HSPF. Complications arise under several circumstances: (1) if a configuration other than the "Individual" one is required, (2) if dealing with impervious area, or (3) if raster landuse data is used instead of polygon data.

To maintain consistency, the "Individual" option has been adopted as the standard for implementing the ArcGIS Timeseries Preprocessing methodology. Though the ArcGIS HSPF Preprocessing methodology can be used to build a new .uci file using the "Grouped" option, no effort has been spent to create or maintain a geospatial representation of this type of model configuration. All of the tools of the ArcGIS Timeseries Preprocessing methodology assume that a .uci file is built using the "Individual" option.

If impervious land is simulated with HSPF, a single feature shown in Figure 4.13 may represent the area of two HSPF Model Elements, one Pervious Land Segment and one Impervious Land Segment. To workaround this problem, two attributes are used to link the GIS representation of model elements to the Elements defined in the .uci file; one for Pervious Land Segments and one for Impervious Land Segments. These attributes are shown in Figure 4.11 and discussed in Section 5.4.

A method for using raster landuse data is presented in Appendix A. Rather than maintaining an explicit polygon feature representation of each type of landuse in each drainage area, the shape of the entire drainage area is retained to give an approximate location for the Land Segment.

4.3.3 HSPFCode – Link between GIS and .uci file

The methodology developed here makes the link between Model Elements in the .uci file and GIS data explicit by storing the same character string in the .uci file and in GIS data. The basic concept behind this character string, called the HSPFCode, is to have a unique, model-wide identifier in both the .uci file and in the geodatabase. Figure 4.11 shows GIS data and a .uci file with the HSPFCode assigned based on the Operation Numbering convention.

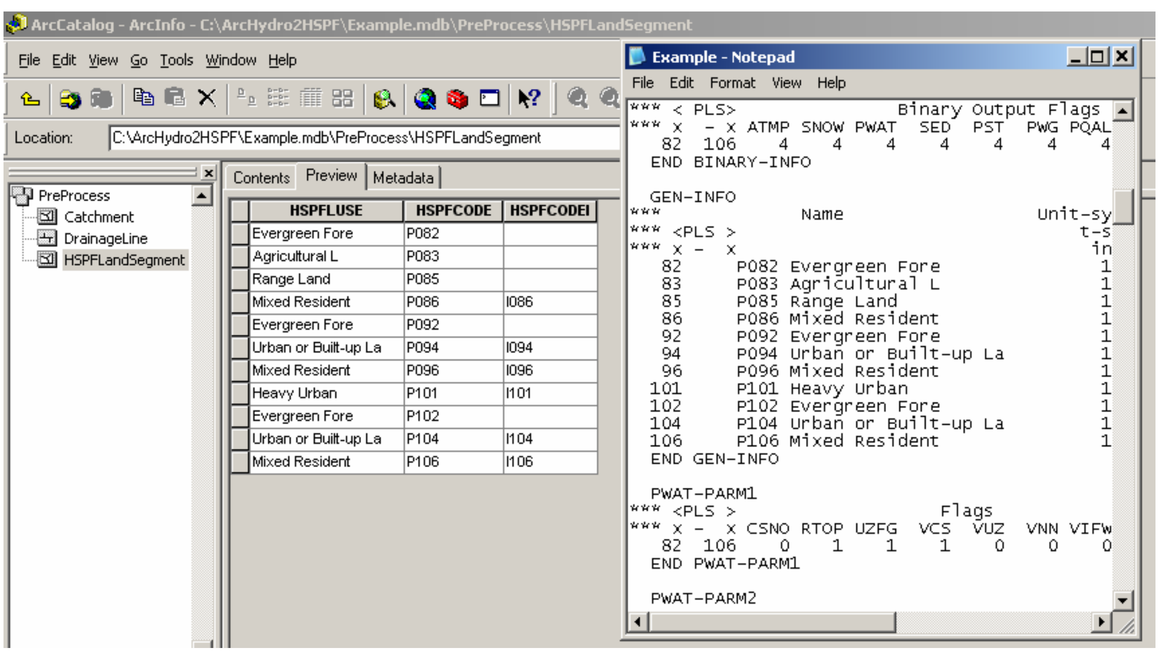


Figure 4.11 HSPFCode: Link between .uci file and GIS data.

Though a tool included in the ArcGIS HSPF Preprocessing system assigns the HSPFCode based on the Operation Numbering convention used by WinHSPF, the concept of the HSPFCode is general and would work with any configuration or Operation Numbering convention. The HSPFCode can be any model-wide, unique identifier that is used to identify Model Elements in both the .uci file and in the geodatabase.

The HSPFCode is used in an implementation of the ArcGIS Timeseries Preprocessing methodology presented in Chapter 5. The .uci file is automatically updated so that Model Elements receive inputs from the appropriate time series datasets based on information stored in the geodatabase. The details of where the HSPFCode is stored and how it is implemented are presented in Section 5.4.

Chapter 5 Procedure of Application

The ArcGIS HSPF Preprocessing system and the ArcGIS Timeseries Preprocessing system have both been implemented in the ArcGIS environment using a structured geodatabase design and ArcToolbox tools. The implementation of the ArcGIS HSPF Preprocessing system includes a database design to store GIS data during the preparation of an HSPF model and a set of tools to perform spatial analysis on GIS data and write the necessary files to build a new model. The implementation of the ArcGIS Timeseries Preprocessing system includes an extended geodatabase design and custom tools to write Arc Hydro time series to .wdm file datasets and update HSPF model files to read from these datasets.

Figure 5.1 illustrates the components of both the ArcGIS HSPF Preprocessing system and the BASINS HSPF Preprocessing system. The ArcGIS HSPF Preprocessing system starts with GIS data from existing tools, specifically, the Arc Hydro terrain processing tools, and performs GIS calculations using new and existing ArcGIS tools. After making the required calculations, custom ArcGIS HSPF Preprocessing tools extract necessary information to intermediate text files for use with WinHSPF, a non-GIS components of BASINS. Custom ArcGIS tools and data structures are also used to prepare precipitation data for HSPF modeling.

The first two sections of this chapter describe the programming tools used in this research and provide instructions for installing tools. The remaining sections give detailed descriptions of the geodatabase structure and tools used to implement the ArcGIS HSPF and Timeseries preprocessing methodologies.

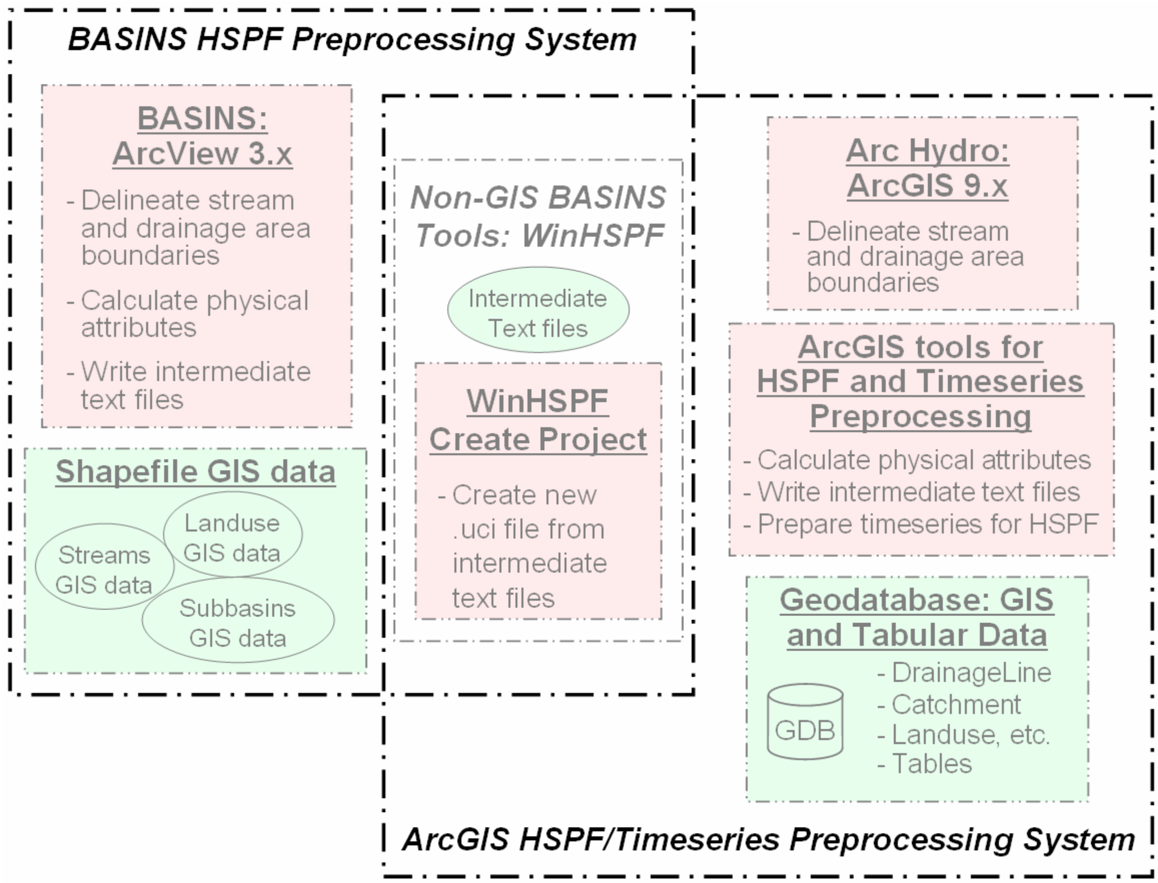


Figure 5.1 BASINS – ArcGIS HSPF/Timeseries Preprocessing comparison.

5.1 PROGRAMMING TOOLS

All of the new tools developed for the ArcGIS HSPF and Timeseries Preprocessing systems are programmed in Microsoft Visual Basic 6. Two separate programming libraries are used to implement the GIS-related and HSPF-related algorithms used in the HSPF and Timeseries Preprocessing systems.

ESRI's ArcObjects are COM-Compliant "Application Programming Interfaces" (API) that are used by software developers to create custom applications within the ArcGIS environment. ArcObjects expose the functionality of ArcGIS software (including ArcMap and ArcToolbox functions) in a coding language so that existing and custom functions can be used to create new tools for specific applications. (ESRI 2005a)

The second API used in the ArcGIS HSPF and Timeseries Preprocessing systems is a set of libraries collectively referred to as the WinHSPF/GenScn/WDMUtil Libraries. These libraries expose the functionality of the WinHSPF/GenScn/WDMUtil programs in a similar way as ESRI's ArcObjects.

5.1.1 ArcObjects and Geoprocessing Functions

A quick, convenient way to implement an algorithm in the ArcGIS environment is through the use of custom Geoprocessing Function Tools (ESRI 2005b). In the ArcGIS 9.0 system, hundreds of built-in Geoprocessing tools are available in the ArcToolbox application that can be accessed in either ArcMap or ArcCatalog. A Geoprocessing Function Tool can be thought of simply as an operation that takes a set of inputs (GIS or non-GIS data) and produces a set of outputs. These tools perform everything from complex spatial analysis such as Zonal Statistics to simple topological operations such as Merge, Clip, or Intersect. Figure 5.2 shows the "Clip" Geoprocessing tool from the ArcToolbox application accessed from ArcCatalog. The tool takes in two GIS datasets and performs the operation described at the right. The location of the resulting dataset is also a required input to the Geoprocessing tool as well as an optional 'Cluster Tolerance' value.

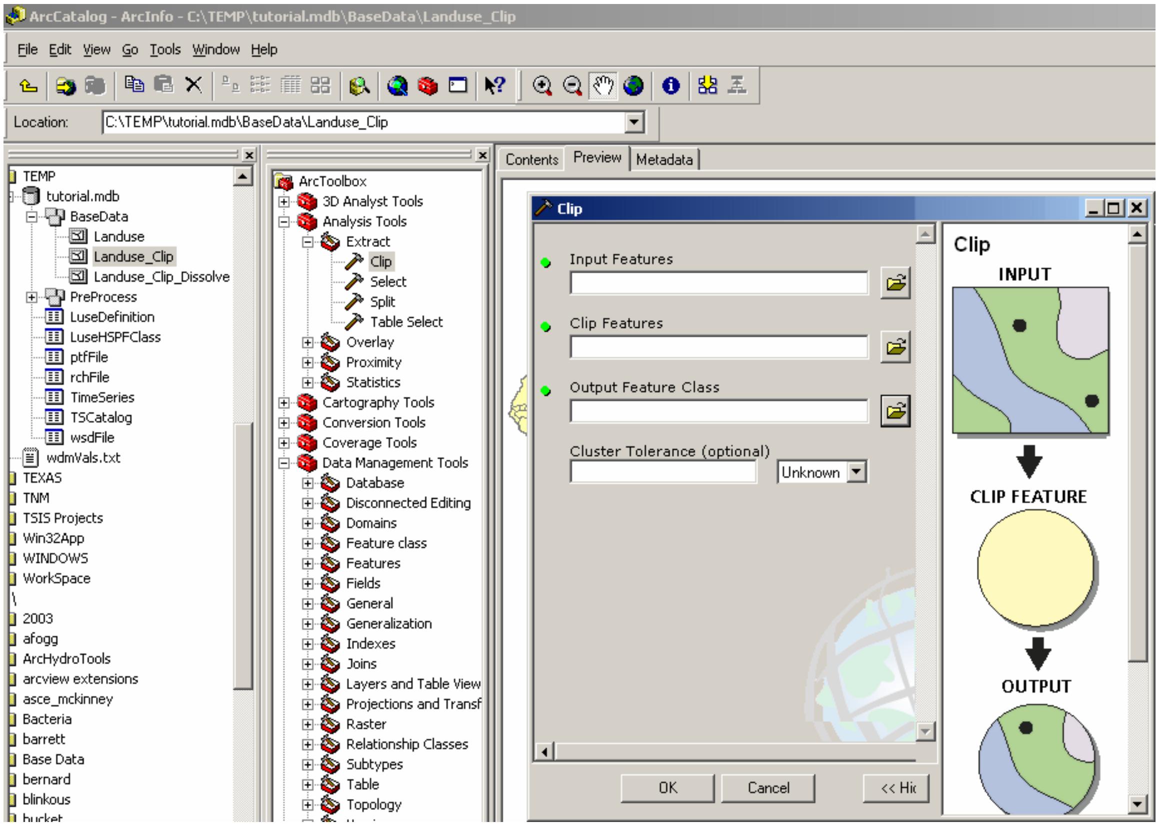


Figure 5.2 ArcGIS Geoprocessing function: 'Clip Features.'

Some tools required to implement the ArcGIS HSPF Preprocessing system were not available in the standard package of tools distributed with ArcGIS. As an example, one component of the ArcGIS HSPF Preprocessing methodology involves writing intermediate text files required by the WinHSPF "Create New Project" tool for building a new .uci file. To implement this process, a custom Geoprocessing tool was created that takes a feature class and table as inputs and writes the required data to the specific format of the intermediate text files.

Custom Geoprocessing tools were developed using the Visual Basic 6 programming language and the ArcObjects API for working with GIS data. Once the programs are compiled, they can be distributed as Digital Link Libraries (DLLs) and used on any Windows machine that has ArcGIS software and libraries installed. The details of each of the geoprocessing tools involved in the ArcGIS HSPF and Timeseries Preprocessing systems are described in Sections 5.3-5.5.

5.1.2 AQUA TERRA WinHSPF Libraries

WinHSPF, GenScn, and WDMUtil are components of the BASINS system that operate independently of the ArcView 3.x GIS component of BASINS. The programs contain many tools for maintaining and manipulating HSPF .uci files and organizing and managing .wdm time series files. They are distributed free of charge from the EPA's BASINS website and AQUA TERRA's WinHSPF site. (EPA 2005b, AQUA TERRA 2005) The WinHSPF/GenScn/WDMUtil programs were developed with the Microsoft Visual Basic 6 programming language using an object-oriented structure so that components used in the programs are also available for use in custom applications.

The WinHSPF interface is designed to expose all the parameters and structure of a .uci file to the user through a Windows interface. In order to accomplish this, the entire .uci file is read into the program's memory using a structure that is related to the .uci file structure, but hidden behind the program interface. All the information from a .uci file resides in the computer's memory within a computer code structure that makes it easy to manage, change, and manipulate parameters associated with each Model Element. The structure of the code used to store the .uci file in the program memory is not specific to the tools available in the WinHSPF software and, consequently, is also available to external tools for performing custom operations on .uci files which are not supported by the WinHSPF program.

One component of the ArcGIS Timeseries Preprocessing methodology involves updating .uci model files to read time series from datasets specified in a geodatabase. To implement this process, a custom Geoprocessing tool was created which first reads the entire .uci file into the program's memory. If the HPSF model has been georeferenced to GIS data, as described in Section 4.3.3, the HSPFCode for each Model Element is used to query the geodatabase and find the time series for the appropriate GIS MetSegment. The computer code storing the .uci file in the computer's memory (the same code that lies behind WinHSPF) is then used to update entries for each Model Element in the .uci file, specifying which time series is to be used in the simulation.

Any computer that has the entire BASINS system or GenScn/WinHSPF/WdmUtil package installed has the libraries needed for the ArcGIS HSPF Preprocessing tools to function correctly. The tool to assign MetSegments to the .uci file requires the latest update to the GenScn package, (Update 4, 4/27/05).

5.1.3 Model Builder and Linking Geoprocessing Tools

One conceptual difference between the ArcGIS HSPF Preprocessing methodology and the BASINS HSPF Preprocessing Methodology is the exposure of many of the individual functions that were hidden behind the user interface in the ArcView 3.x system. While the ArcView 3.x system combined the "Subbasins" shapefile with the "hspfusgs.dbf" table and wrote the intermediate text files with a single click of a button, the ArcGIS system uses several new and existing ArcToolbox tools to provide the same functionality. Though this is not as simple as a one-button tool, it makes the system more general, giving advanced users access to individual parts of the system.

If a set of processes within the ArcGIS HSPF Preprocessing system are performed repeatedly, individual tools can be linked together using the ArcGIS ModelBuilder application to streamline the process into a one-button-type tool. The ModelBuilder application is a sort of "visual programming" interface for ArcGIS Geoprocessing functions (ESRI 2005c). Tools from ArcToolbox can be linked together by dragging them and the associated data onto a form. Output from one tool can be used as input to another tool so that several complex functions can be performed with a one-click model. Figure 5.3 shows a ModelBuilder model in which 1) one feature class is used to clip out part of a larger dataset; 2) the resulting feature class is dissolved based on an attribute;

and 3) the resulting feature class is intersected with another feature class to produce the final result.

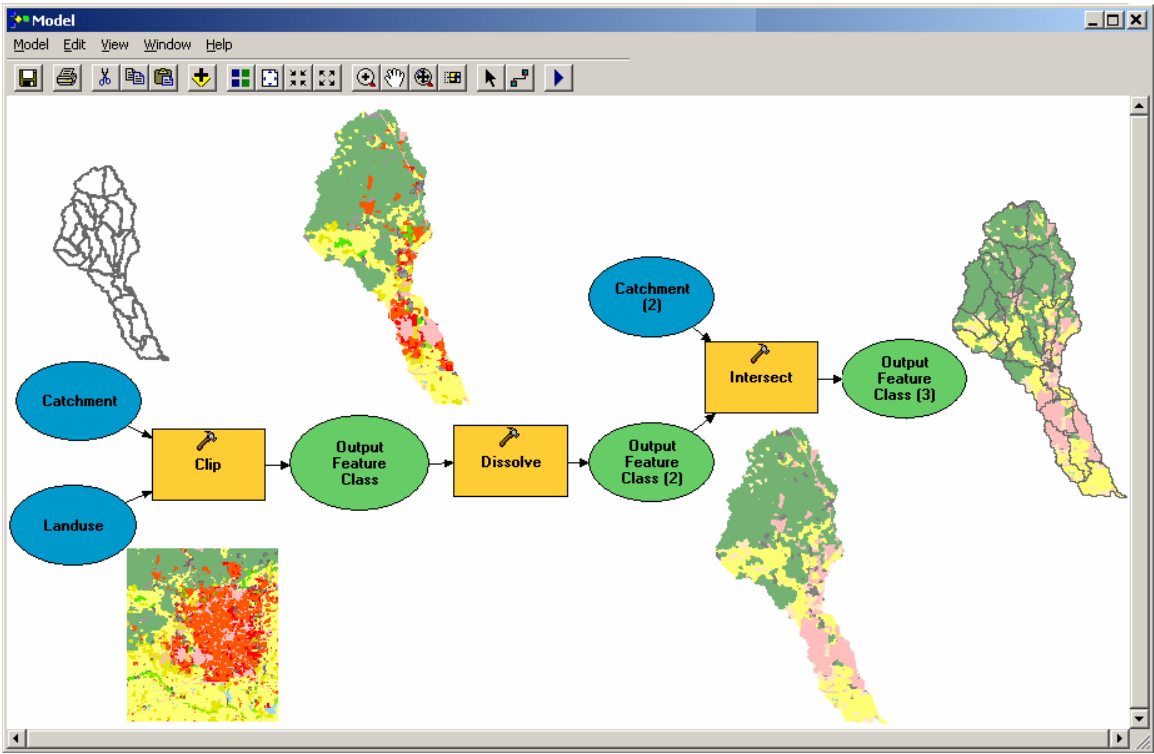


Figure 5.3 ArcGIS ModelBuilder example.

In addition to providing a way of combining many functions into a single tool, or model, the ModelBuilder application provides an intuitive, visual way of presenting the concepts in a complex process. The ArcGIS HSPF Preprocessing Methodology is, at its core, a framework for using ArcGIS tools to prepare data for HSPF modeling. Though the entire system could be implemented in a single ModelBuilder model, it does not have to be. The ArcGIS HSPF and Timeseries Preprocessing systems can be implemented by linking tools together in ModelBuilder or, equivalently, by performing each of the functions in sequence using the individual ArcToolbox tools.

5.2 TOOL PRESENTATION

The tools used to implement the ArcGIS HSPF and Timeseries Preprocessing methodologies are all Geoprocessing Functions, which are accessed through ESRI's ArcToolbox environment. Some of the tools use standard ESRI functions and others are custom tools developed specifically for the purposes of this research. This section introduces the tools and provides an explanation of how to install them.

5.2.1 Installation Instructions

The ArcGIS HSPF and Timeseries Preprocessing tools are contained in the 'crwrHSPFTools.zip' file located on Appendix E (the zip file attached to this thesis). The tools are installed by extracting the contents of this zip file to the working directory ('C:\' directory is preferable) on a computer and running the '_Install.bat' file to register the tools with the computer and with ESRI software.

5.2.2 File Descriptions

The 'crwrHSPFTools.zip' file contains a primary directory '\ArcHydro2HSPF' and several subdirectories. The 'bin' directory contains the Digital Linked Library (.dll) file, crwrHSPFTools.dll, which contains the code for custom ArcToolbox functions. The 'bin' directory also contains template files used by several of the tools. The 'TemplateTables' directory contains examples of the tables used in the ArcGIS HSPF and Timeseries Preprocessing methodologies.

The 'Tutorial' directory contains sample data to be used in applying the ArcGIS HSPF Preprocessing methodology, and the 'xml' directory contains .xml files defining the structure of the ArcGIS HSPF Preprocessing geodatabase. These .xml files are intended to be used with the ESRI Geodatabase Designer2 extension, which can be downloaded from: http://arcscripts.esri.com/details.asp?dbid=13484 or found on Appendix E (also includes tutorial on preparing Geodatabase).

Figure 5.4 shows the 'ArcHydro2HSPF' directory in ArcCatalog.

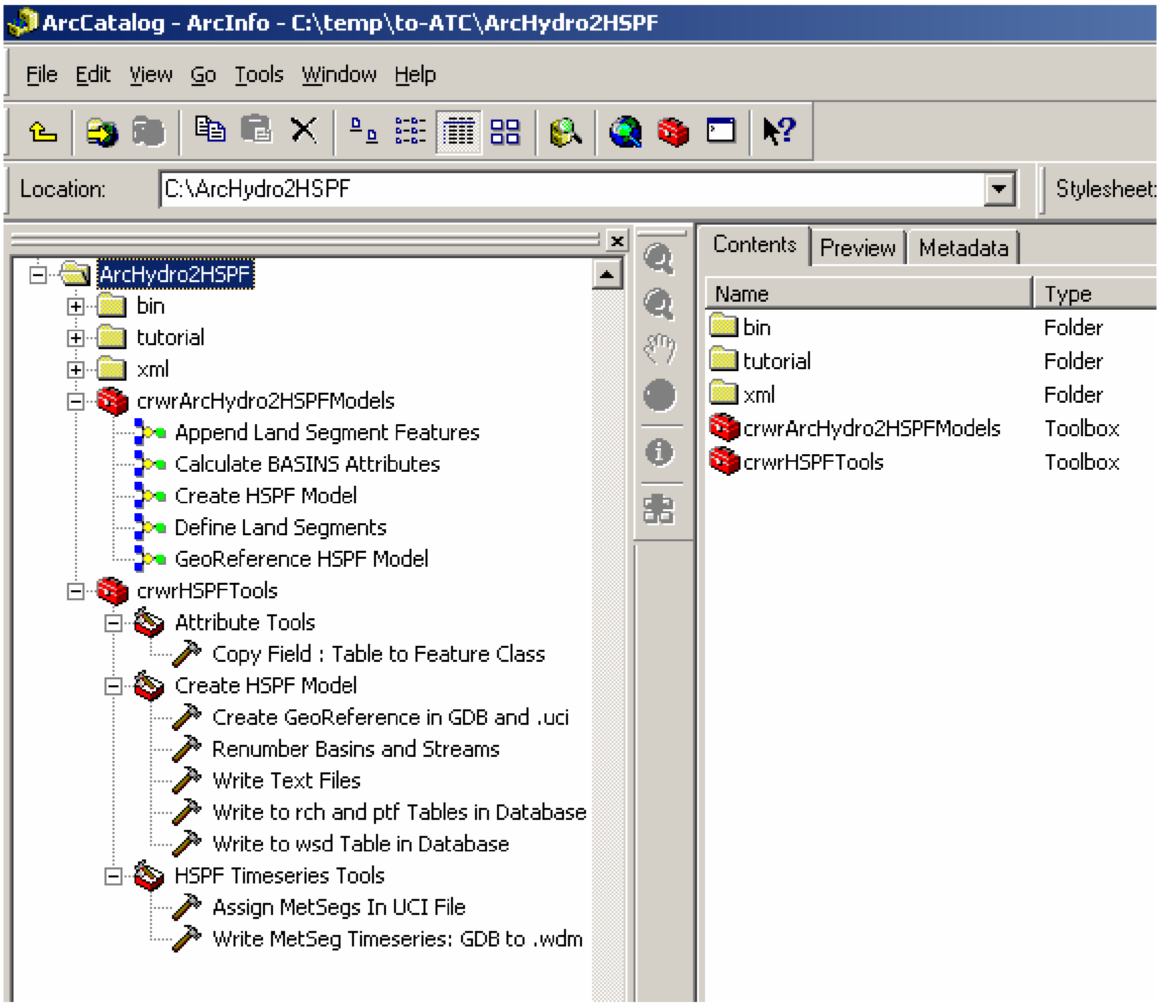


Figure 5.4 'ArcHydro2HSPF' directory.

Two ArcToolbox toolsets reside in the 'ArcHydro2HSPF' directory. These toolboxes contain the tools and models used to implement the ArcGIS HSPF Preprocessing methodology. The 'crwrHSPFTools' toolbox contains the custom ArcGIS tools developed for this research. The tools are designed to be implemented in the ArcCatalog environment, using feature classes and not feature layers in ArcMap. The 'crwrHSPFModels' toolbox contains a set of ModelBuilder models that link together standard and custom Geoprocessing Functions to streamline the application of the tools. Sections 5.3.2, 5.4.1, and 5.5.2 outline each of the tools in detail and describe how they are applied.

5.3 ARCGIS HSPF PREPROCESSING IMPLEMENTATION

The the ArcGIS HSPF Preprocessing system includes:

1) A database design to organize and store GIS data during the preparation of an HSPF model
2) A set of existing and custom tools that use data in the geodatabase to perform spatial analysis and write the model files.

Though the tools are intended to be used with data stored in a manner consistent with the structured geodatabase design, they are not hard-coded to it. This means that data stored in a format with different filenames and attribute names can be used with the tools as long as all the necessary information is present. An attempt is made to maintain consistency with the BASINS HSPF Preprocessing system so that data developed with the BASINS 3.x system can still be easily used in the ArcGIS environment. As such, some attribute names from the BASINS 3.x shapefiles are used to store the same information in the ArcGIS geodatabase design.

In the following sections, custom Geoprocessing functions are presented in detail, and ESRI help files should be consulted for documentation of standarad ESRI Geoprocessing functions.

5.3.1 Geodatabase Design

Overview

The geodatabase design to support the ArcGIS HSPF Preprocessing system contains four feature classes organized into two feature datasets, “BaseData” and

“PreProcess.” Two tables, LuseDefinition and LuseHSPFClass, are used directly in the preprocessing application, and three other tables are used to transfer data to intermediate text files. Figure 5.5 illustrates the geodatabase with the two feature datasets and five tables.

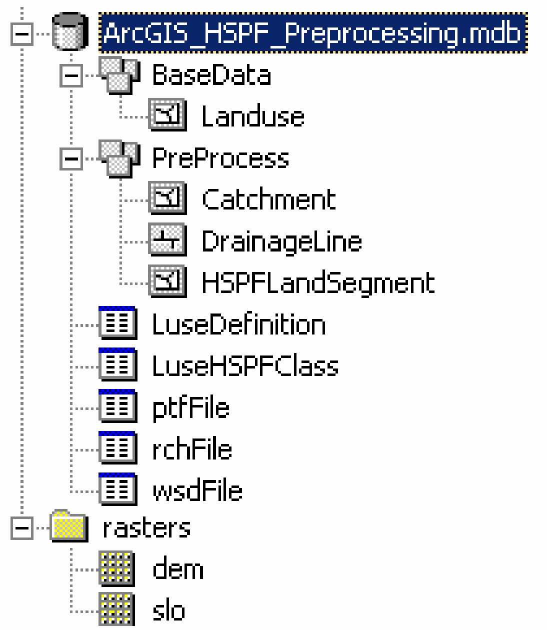


Figure 5.5 ArcGIS HSPF Preprocessing geodatabase structure.

A more detailed illustration of the geodatabase design is presented in Appendix C. The drawing contains detailed specifications for the attributes of each geodatabase object and presents the relationships that are used to implement the ArcGIS HSPF Preprocessing system. These relationships are not explicit geodatabase classes, but are used by the ModelBuilder models and custom Geoprocessing tools to transfer information amongst GIS data.

The only feature class in the “BaseData” feature dataset is the “Landuse” feature class. Other data describing the landscape such as soils, NHD streams, or other data can be stored in this feature dataset, but only the “Landuse” feature class is used by the ArcGIS HSPF Preprocessing system. The “PreProcess” feature dataset contains two feature classes, “Catchment” and “DrainageLine” which must be present and populated to use the tools, and a third feature class, “HSPFLandSegment” that is populated using results from the tools. This feature class was presented in Figure 4.10 and is intended to aid in defining a geospatial representation of the Land Segments simulated by an HSPF model.

Tables used in HSPF Preprocessing reside outside the feature datasets but still within the geodatabase. Two of the tables, “LuseDefinition” and “LuseHSPFClass,” are associated with defining landuse types to be simulated by with an HSPF model and must be present and populated to use the tools. The other three tables are populated using ArcGIS HSPF Preprocessing tools.

Two raster datasets are used to calculate slopes for Land and River Segments in the ArcGIS HSPF Preprocessing system. These rasters are shown in a separate file folder outside the geodatabase in Figure 5.5, but they could equivalently be stored in a raster dataset. Raster Datasets are a geodatabase structure new in ArcGIS 9 similar to feature datasets, but used to store raster datasets. The ArcGIS HSPF Preprocessing system uses many standard ArcGIS tools which can access data stored in any location, so the specifics of where raster datasets are stored is not critical to the system.

The following sections present the geodatabase feature classes and tables as well as their attributes in detail. Appendix C contains a large drawing of the geodatabase and may aid in understanding relationships amongst datasets.

DrainageLine and Catchment Feature Classes

“Catchment” contains the drainage areas for each of the Reach Segments to be modeled in HSPF, and “DrainageLine” contains the River Reach Segments. The names “DrainageLine” and “Catchment” are used for the river segments and associated drainage areas in the ArcGIS HSPF Preprocessing system as opposed to “Subbasins” and “Streams” from the BASINS HSPF Preprocessing system. Those familiar with the

BASINS system will notice that there is no equivalent for the “Outlets” shapefile in the ArcGIS HSPF Preprocessing geodatabase. The “Outlets” shapefile is only used by the BASINS system if point sources are being considered, and this functionality is not supported by the ArcGIS system.

The Arc Hydro Terrain PreProcessing tools create two feature classes called “DrainageLine” and “Catchment” which have a minimal set of attributes defining the connectivity between river segments and their associated drainage area as well as network connectivity (Maidment 2002). Three of these attributes are used explicitly by the ArcGIS HSPF Preprocessing methodology. “HydroID” is used as a unique identifier for every feature in an Arc Hydro geodatabase and is present on both the DrianageLine and Catchment feature classes. “DrainID” of DrainageLine features contains the “HydroID” of the corresponding Catchment. “NextDownID” on Catchment features contains the “HydroID” of the next downstream Catchment.

There is a one-to-one (and only one) relationship between Catchments and DrainageLines produced with the ArcHydro Terrain PreProcessing tools. Each Catchment must have a single corresponding feature in the DrainageLine feature class which represents the river segment draining the Catchment.

Figure 5.6 provides a summary of the attributes of the “Catchment” and “DrainageLine” feature classes in the ArcGIS HSPF Preprocessing geodatabase. The “DrainageLine” and “Catchment” feature classes in the ArcGIS HSPF Preprocessing geodatabase contain the attributes from the associated Arc Hydro feature classes as well as some attributes from the BASINS HSPF Preprocessing system. The primary function of the ArcGIS HSPF Preprocessing geodatabase is to develop inputs to the HSPF model, as opposed to the BASINS system in which GIS data is used for many other purposes as well. Only the attributes from the BASINS system that are required for building a new HSPF model are included in the feature classes of the ArcGIS HSPF Preprocessing geodatabase. A detailed drawing of the geodabase design is presented in Appendix C, and a brief description of the attributes for each element in the database will be given here.

- **Catchment feature class**
  - *Arc Hydro Attributes*
    - HydroID
    - HydroCode
    - DrainID
    - AreaSqKm
    - JunctionID
    - NextDownID
  - *BASINS Attributes*
    - Subbasin
    - Slo1
  - *ArcGIS Attributes*
    - Shape_Length
    - Shape_Area

- **DrainageLine feature class**
  - *Arc Hydro Attributes*
    - HydroID
    - HydroCode
    - DrainID
  - *BASINS Attributes*
    - Subbasin
    - Subbasinr
    - Slo2
    - Wid2
    - Dep2
    - MinEl
    - MaxEl
  - *ArcGIS Attributes*
    - Shape_Length

Figure 5.6 ArcGIS HSPF Preprocessing feature class attributes: Catchment, DrainageLine.

The ArcHydro attributes required by the ArcGIS HSPF Preprocessing system are "HydroID" and "NextDownID" from the Catchment feature class and "DrainID" from the DrainageLine feature class. Three of the BASINS attributes, "Subbasin" and

“Subbasinr” from DrainageLine and “Subbasin” from Catchment, contain almost identical information as the Arc Hydro attributes. The only difference is that downstream information is stored on the DrainageLine instead of the Catchments. Though it is redundant to store this information twice, an attempt is made in the ArcGIS system to maintain consistency with the naming conventions of the BASINS system. A tool is available in the ArcGIS HSPF Preprocessing system to populate the “Subbasin” and “Subbasinr” attributes or, alternatively, data from the BASINS system can be loaded directly into the Catchment and DrainageLine feature classes.

The BASINS attributes for the slope of the land surface (Slo1) and river slope (Slo2) are calculated in the ArcGIS HSPF Preprocessing system with tools in ArcToolbox using Digital Elevation Model and slope raster data. The ArcGIS system does not have tools to calculate the width and depth of River Segments, but these attributes are necessary to create FTABLES in the new .uci file. If estimates of stream width and slope are available from another source, they can be manually entered into the “Wid2” and “Dep2” attributes of the DrainageLine feature class. If no estimates for river geometry are available, the ArcGIS HSPF Preprocessing tools will assume default values when creating a new .uci file.

It is important to note that the ArcGIS HSPF Preprocessing system assumes that each feature in the DrainageLine feature class corresponds to a River Segment Model Element to be simulated with HSPF. In addition, it is assumed that there is a one-to-one (and only one) relationship between Catchment features and DrainageLine features. This type of data structure is commonly the result of terrain processing tools such as ArcHydro, but is not common in other data sources such as NHD Flowlines or HUC boundaries. If data is used in the ArcGIS HSPF Preprocessing system that is not the result of the Arc Hydro Terrain Processing tools, it should be carefully checked to make sure it meets these assumptions. The DrainageLine feature class should contain a single feature for each River Segment to be simulated in HSPF, and the Catchment feature class should have one and only one drainage area for each DrainageLine feature.

Landuse and Associated Tables

Two of the five tables in the HSPF Preprocessing geodatabase are directly associated with managing the landuse types to be simulated by the HSPF model. The "LuseDefinition" table contains the information to transfer landuse categories from the Landuse feature class to new categories for use in HSPF modeling. "LuseHSPFClass" contains the information to define Impervious Land Segments and to define the Operation Numbering convention to be used in the new HSPF .uci file.

- **Landuse feature class**
  - *LULC Attributes*
    - LUCODE
    - LEVEL2
  - *ArcGIS Attributes*
    - Shape_Length
    - Shape_Area
- **LuseDefinition table**
  - LUCODE
  - HSPFLUSE
- **LuseHSPFClass table**
  - HSPFLUSE
  - PLuseID
  - PercentImp
  - HSPFLUSEI
  - ILuseID

Figure 5.7 ArcGIS HSPF Preprocessing landuse attributes.

Figure 5.7 provides a summary of the attributes of the "Landuse" feature class, "LuseDefinition," and "LuseHSPFClass" tables in the ArcGIS HSPF Preprocessing geodatabase. The only attribute from the Landuse feature class used in the ArcGIS HSPF Preprocessing system is "LUCODE." This attribute contains a unique integer value with a corresponding value in the LuseDefinition table. The "LEVEL2" attribute name is specific to the Land Use Land Cover dataset distributed by the USGS. It contains a text description of the Level 2 Anderson Classification category, but is not used by the ArcGIS HSPF Preprocessing system.

The "LUCODE" attribute in the LuseDefinition table corresponds to the "LUCODE" attribute of the landuse feature class. Each record in the landuse feature class is mapped to an HSPF landuse type (to be simulated by the HSPF model) using the "LUCODE" attribute to join the two tables together. A record must exist in the LuseDefinion table for every unique value in the Landuse feature class. "HSPFLUSE" is a 20-character string describing the categories of land to be modeled with HSPF. There should be no more than nine unique types of HSPFLUSE in the LuseDefinition table. Each of these (up to nine) "HSPFLUSE" types introduced in the LuseDefinition table must have a corresponding record in the LuseHSPFClass table.

In the LuseHSPFClass table, a record with a matching HSPFLUSE attribute must be present for each HSPFLUSE type introduced in the LuseDefinition table. The "PLuseID" attribute defines the third digit used in the Operation Numbering convention for Pervious Land Segments and the "ILuseID" attribute is the third digit for Impervious Land Segments. For every record (landuse type) that has a "PercentImp" greater than zero, the "HSPFLUSEI" and "ILuseID" must be populated as well. The "HSPFLUSEI" attribute is also a 20-character string defining the Impervious Land category that the landuse type belongs to. A unique "HSPFLUSEI" can be assigned for each landuse type, or more than one type can be joined together by assigning the same "HSPFLUSEI" and "ILuseID" attribute.

The Operation Numbering convention used by the BASINS system was presented in Section 3.4. Each Land Segment Operation in an HSPF model that simulates the same type of landuse is assigned the same third digit in its Operation Number. This Operation Numbering scheme is inherently limited to 10 types of landuse (for a decimal numbering system), however, the WinHSPF "Create Project" tools are only capable of using nine unique types of landuse.

If the “Individual” option in the WinHSPF “Create Project” tools is used to create a unique set of landuse Operations for each River Segment simulated by the model, this type of model configuration could result in hundreds of Land Segment Operations for a large river network. Care should be taken to ensure that the HSPF limitation of 500 Model Operations in a single model run is not exceeded.

HSPFLandSegment and Intermediate Geodatabase Tables

The HSPFLandSegment feature class contains only the attributes required for defining Land Segments for HSPF modeling. This feature class can be populated using the results of some of the tools presented below. The feature class will contain an explicit spatial representation of Model Elements. The intermediate geodatabase tables, wsdTable, rchTable, and ptfTable are used to store the information for the intermediate text files in the geodatabase. They are populated from GIS data using tools presented below and, once populated, contain all the data required to write the intermediate text files for building a new HSPF .uci file.

5.3.2 ArcToolbox Implementation

The ArcGIS HSPF Preprocessing methodology uses as many existing tools as possible to bridge the gap between ArcGIS and HSPF. The methodology presented here assumes that existing Arc Hydro, BASINS, or other terrain processing tools have been used to produce the “Catchment” and “DrainageLine” feature classes. Three basic tasks are accomplished in the ArcGIS environment to prepare data for input to the HSPF model:

1) calculate physically-based attributes required for HSPF model creation
2) define Land Segments (in GIS) to be simulated by HSPF model
3) write intermediate text files for use with WinHSPF “Create Project” tool

Once these tasks have been accomplished, WinHSPF can be used to build a new .uci file outside the GIS environment.

The tasks presented above can be implemented either with the ModelBuilder models included with the ArcGIS HSPF Preprocessing tools, or by equivalently applying each Geoprocessing function included in the model in order. The remainder of this section presents a methodology for applying the tools using the ModelBuilder models. Custom ArcGIS tools developed specifically for this research are described in detail, and ModelBuilder models which use standard ArcGIS tools are described only briefly.

Load Data and Renumber Basins and Streams

Data resulting from the application of the Arc Hydro Terrain PreProcessing tools or from another source should be loaded into the ArcGIS HSPF Preprocessing geodatabase before applying the tools. A detailed description of how to create a geodatabase with the necessary structure and load data into it is given in a tutorial in Appendix E.

Before calculating physically-based parameters for HSPF model development, a tool is included in the ArcGIS HSPF Preprocessing tools that aids in numbering GIS features in a way more consistent with the BASINS Preprocessing methodology. A custom tool was developed for the ArcGIS HSPF Preprocessing system to populate the "Subbasin" and "Subbasinr" BASINS attributes of DrainageLine and Catchment. The same network information is stored in the Arc Hydro attributes "HydroID" and "NextDownID," but its format is slightly different. Downstream information is stored on the River Segment features in the BASINS attributes and on the Catchment features in the Arc Hydro attributes.

Figure 5.8 shows the user-interface to a tool designed to renumber Catchments and DrainageLines by populating the "Subbasin" and "Subbasinr" fields using information in the "HydroID," "NextDownID," and "DrainID" fields.

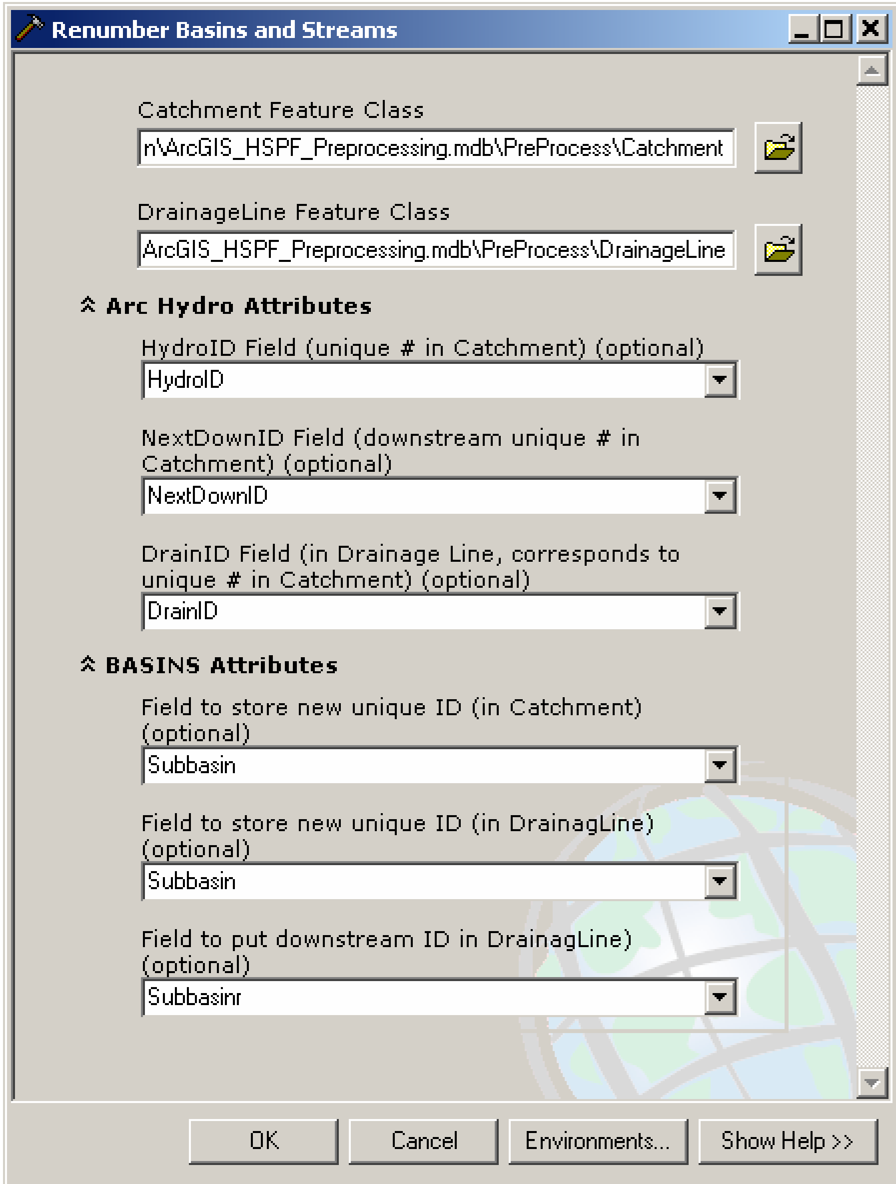


Figure 5.8 Custom ArcGIS Geoprocessing function: Renumber Basins and Streams.

The tool takes two feature classes (Catchment and DrainageLine) as inputs and allows the user to specify the fields to be used in the function. The field names default to values from the ArcGIS HSPF Preprocessing geodatabase but can be changed by the user to match data in a different format. A recursive program is used to trace upstream from the most downstream Catchment and renumber the features (in the “Subbasin” and

“Subbasinr” fields) from one to the number of Catchment features. The most downstream feature is found as the feature with a negative “NextDownID,” typically “-1.” The tool assumes that there is only one outlet from the river network contained in the DrainageLine feature class, and only one feature should have a negative “NextDownID.” The tool does not use a well-defined methodology for renumbering streams except that they are numbered from 1 to n, and no stream segment is assigned a number larger than any stream segment downstream of it.

In addition to maintaining consistency with the BASINS HSPF Preprocessing system attributes, there is another reason for transferring the ArcHydro network attributes to the BASINS network attributes. The “Subbasin” attribute is used as the first two digits of the Operation Number for the convention used by the “Individual” option of the WinHSPF “Create Project” tool. HydroID’s often become larger than two digits in a large Arc Hydro geodatabase, and renumbering the Catchments and DrainageLines starting from one ensures that the “Subbasin” attributes will be smaller than three digits.

It should be noted that the Catchments and DrainageLines do not have to be numbered starting at one. The only requirement for subsequent tools is that downstream information is accurately stored in the “Subbasin” and “Subbasinr” attributes of the Catchment and DrainageLine feature classes. This includes having “-1” for the “Subbasinr” attribute of the most downstream River Segment feature. If the “Individual” option is used to create a new .uci file using WinHSPF, the “Subbasin” attributes must be two digits.

One limitation of the tool to renumber basins and streams is that there must be one ‘most downstream’ river segment. If two features in the DrainageLine feature class have a ‘NextDownID’ of ‘-1’, the tool will not work properly.

Calculate physically-based (slope) attributes

Basic ArcGIS tools available in ArcToolbox are used by the ArcGIS HSPF Preprocessing system to calculate the land surface slope and river slope. Figure 5.9

illustrates a ModelBuilder model that can be used to calculate required BASINS Attributes. Existing tools are linked together into a ModelBuilder model to make the process more efficient. The only new tool used in the slope calculations of Land and River segments, is a tool to “Copy Field: Table to Feature Class.” This tool was developed only to make the ModelBuilder implementation more efficient and is presented in detail at the end of this section.

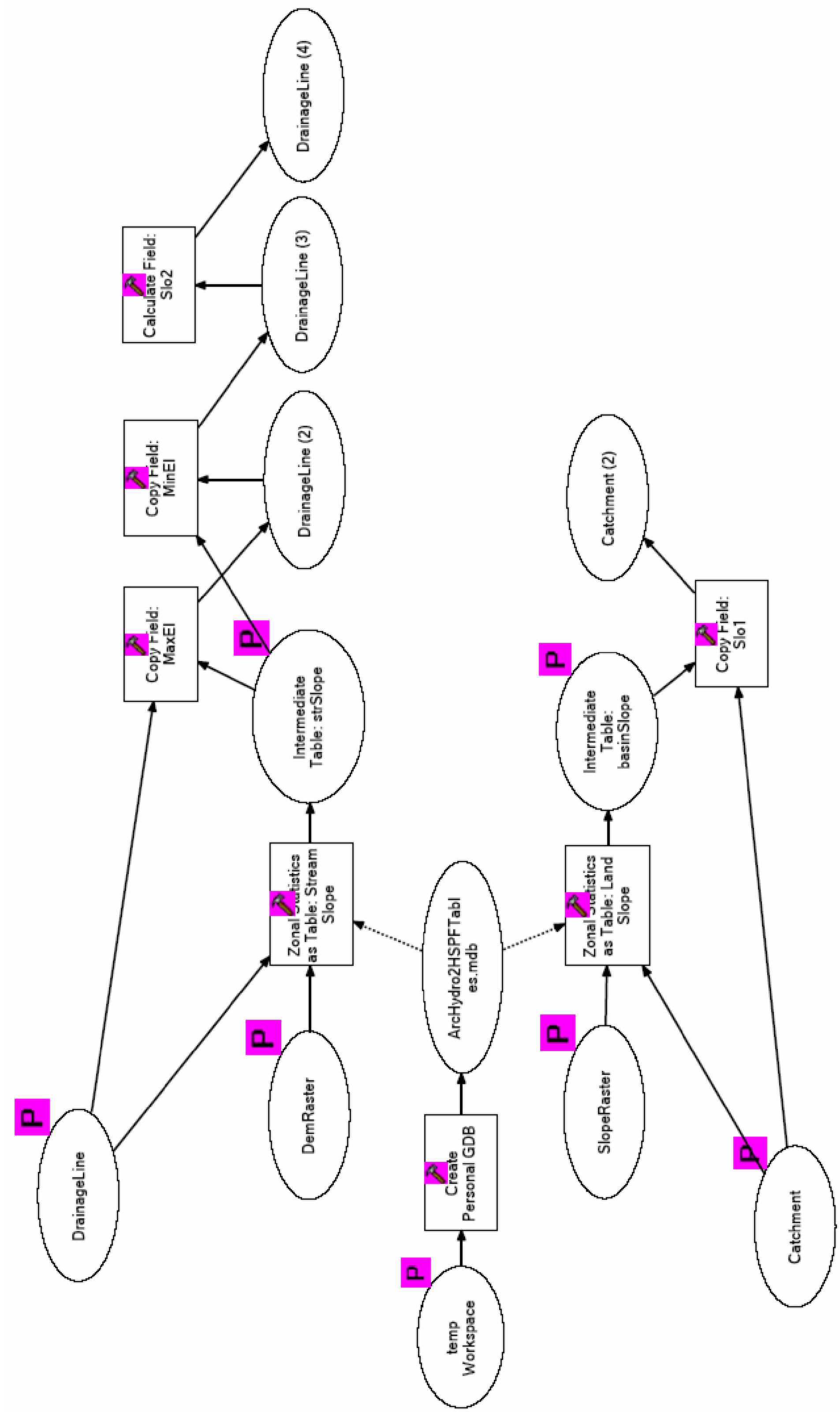


Figure 5.9 ArcGIS ModelBuilder model: Calculate Slope Attributes.

Each of the tools in the model shown in Figure 5.9 are illustrated as rectangle shapes, and the data on which the tools operates (input parameters or tool outputs) are shown as oval shapes. Model input parameters, denoted by a “P,” are used to simplify the model and make it accessible to other input datasets. Primary input parameters to the model include the DrainageLine and Catchment feature classes, as well as DEM and slope rasters. Other parameters are required to define a temporary workspace to store intermediate tables and the paths of those tables.

The upper portion of Figure 5.9 shows the tools used to calculate the slope of the River Segments to be modeled in HSPF. The first tool, “Zonal Statistics as Table: Stream Slope,” takes the DrainageLine feature class and the DEM raster as the primary inputs and calculates statistics for each feature (Max, Min, Mean) and stores the result in a table location also specified as an input. The next two tools, “Copy Field: MaxEl” and “Copy Field: MinEl,” copy the maximum and minimum elevations for each river segment to the appropriate fields in the DrainageLine feature class, “MaxEl” and “MinEl.” The final tool calculates the slope of each river segment by subtracting the minimum elevation from the maximum elevation and dividing by the river length. The DEM raster used in the zonal statistics tool must have the same vertical scale as the linear unit of the DrainageLine feature class projection or a conversion factor is necessary.

The lower portion of Figure 5.9 contains the tools used to calculate the average slope of each feature in the Catchment feature class. Similar to the DrainageLine slope process, the first tool, “Zonal Statistics as Table: Catchment Slope,” takes the Catchment feature class and the slope raster as the primary inputs and the location of the output table as another input. The next tool, “Copy Field: Slo1,” copies the mean slope for each catchment to the appropriate field in the Catchment feature class, “Slo1.”

One additional tool is included which creates a geodatabase workspace to store the intermediate tables. Though this is not necessary, storing the resulting tables from the

"Zonal Statistics as Table" tool in a database was found to make attribute names more predictable.

Figure 5.10 shows the simple User Interface to the model shown in Figure 5.10.

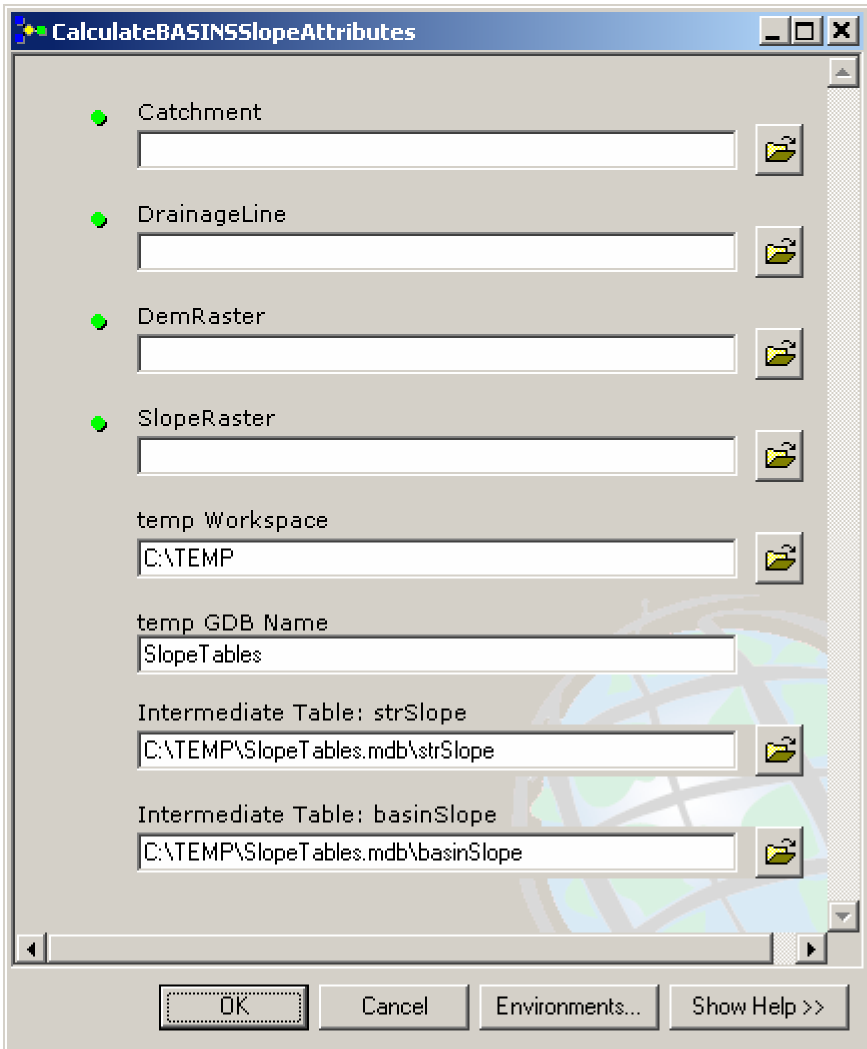


Figure 5.10 ArcGIS ModelBuilder model: Simple user interface.

Model Builder models can be accessed in of two different ways. The "Open" option (or double clicking) opens the interface shown in Figure 5.10 giving access only to the model input parameters. This interface can be used with well-developed models to rapidly apply the model to different datasets. The "Edit" option (right click – choose edit) opens the visual form for the model shown in Figure 5.9. When designing a model or when a model produces errors, it is best to run the model from the "Edit" form to see which tool or dataset within the model is causing the error.

Define HSPF Land Segments

Basic ArcGIS tools available in ArcToolbox are also used by the ArcGIS HSPF Preprocessing system to define HSPF Land Segments using a feature class describing landuse and the Catchments (or drainage area boundaries for the River Segments). A methodology for using raster landuse data is presented in Appendix A. Several tools and datasets are combined into a ModelBuilder model to streamline the geospatial processing required to define the Land Segments for HSPF modeling. Figure 5.11 illustrates a ModelBuilder model that can be used to define HPSF Land Segments from a landuse feature class and a drainage area feature class.

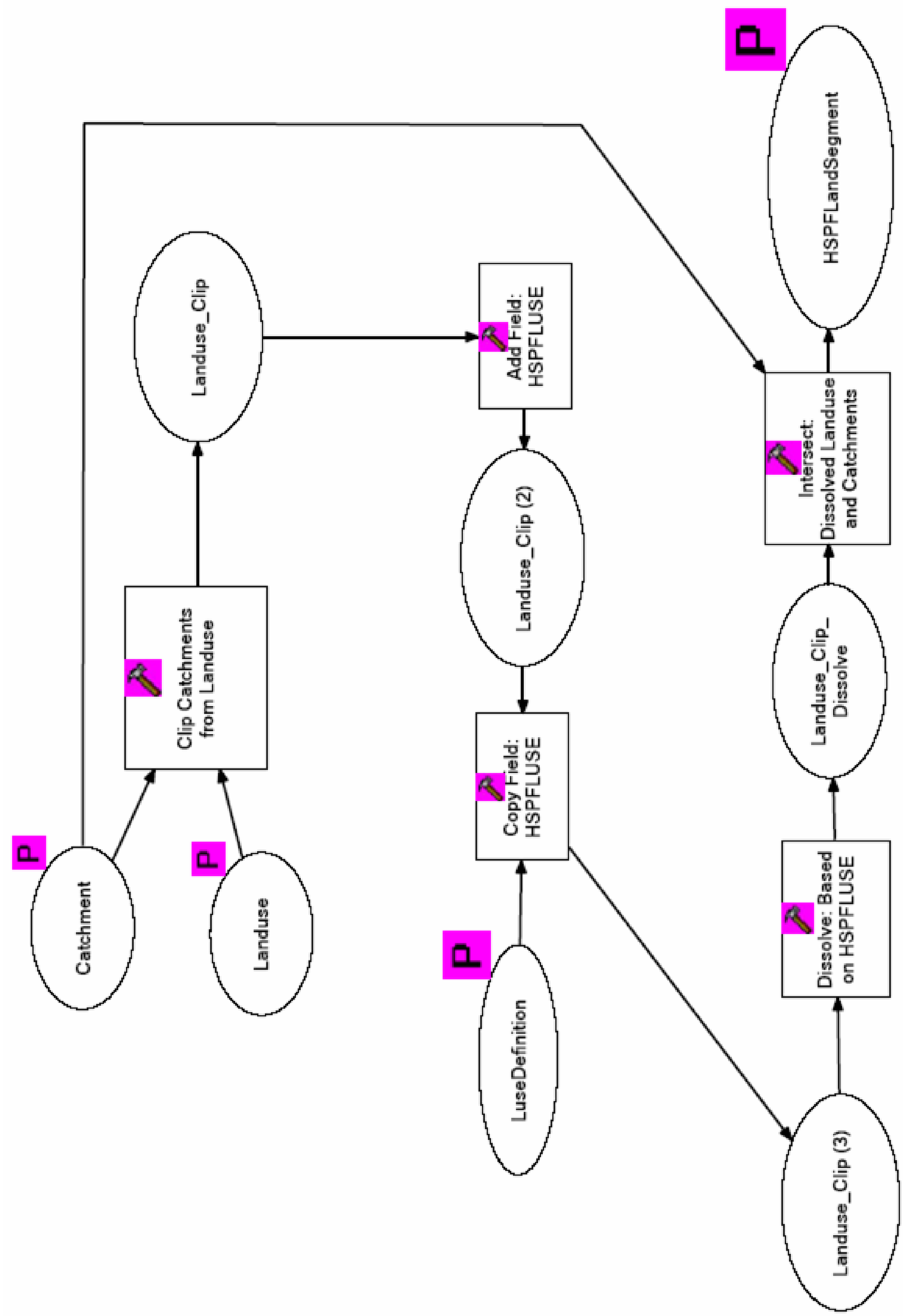


Figure 5.11 ArcGIS ModelBuilder model: Define HSPF Land Segments.

Primary input parameters to the model include the Catchment and Landuse feature classes and the LuseDefinition table used to redefine landuse categories for HSPF modeling. The model takes the output location of the "HSPFLandSegment" feature class as its final input parameter. The first tool in the model "Clip Catchments from Landuse," uses the drainage area boundaries to clip the area of interest out of a potentially large landuse dataset. The second tool "Add Field: HSPFLUSE," adds a field to the resulting feature class that will be used to define the HSPF landuse categories. "Copy Field: HPSFLUSE" copies the HSPF landuse category from the LuseDefinition table to the clipped landuse feature class based on a unique identifier in both the table and landuse data. Another ArcToolbox tool is then used to "Dissolve: Based on HSPFLUSE." The resulting feature class contains a single feature for each HSPF landuse category in the entire watershed. The final process "Intersect: Dissolved Landuse and Catchments," intersects the HSPF landuse feature class with the catchment boundaries to create a feature class that has a set of HSPF landuse features for each drainage area.

The process accomplished by this model could be implemented without the first step of clipping to the catchment boundaries. Output from the final intersection operation would include only the landuse areas inside the Catchment boundaries whether or not the area was clipped from a larger dataset at the beginning. The process of copying the HSPFLUSE field to every feature of an extremely large landuse dataset could be extremely time consuming, so the area of interest is clipped out first.

Defining Impervious Landsegments

In order to use the WinHSPF tools to build a new HSPF .uci file, information from GIS data must be extracted to the intermediate text files used by WinHSPF. Four intermediate files are required to use the WinHSPF "Create Project" tool. The format and content of these text files is detailed in the WinHSPF User's Manual. (Duda 2001)

GIS data for streams can be used directly to extract the data necessary to define HSPF River Segments, however, defining Land segments is more complex. The .wsd file is used to define pervious and impervious land segments for HSPF modeling. Each line of the .wsd file specifies

1) the HSPF landuse category to be simulated by the land segment
2) whether the land segment is pervious or impervious
3) which river segment the land contributes to
4) area of specified landuse type contributing to specified river segment
5) slope of the land segment
6) assumed overland flow distance

All of the information for the .wsd file is present in the ArcGIS HSPF Preprocessing geodatabase after data is prepared using the models described above for "Calculating BASINS Slope Attributes" and "Defining HSPF Land Segments" or an equivalent sequence of Geoprocessing tools. HSPFLandSegment, the result of the "Defining HSPF Land Segments" model presented above, contains the area of each type of landuse contributing to each River Segment as well as the average slope of the drainage area it falls within. Information about impervious area for each HSPF landuse category is stored in the HSPFLuseClass table.

Custom Geoprocessing tools combine information from the HSPFLandSegment feature class and the HSPFLuseClass table to extract information from the geodatabase and write the intermediate text files required for the WinHSPF "Create Project" tool. Three Geoprocessing tools are used in the process of writing the intermediate text files used by WinHSPF to create a new HSPF .uci file. The first two of the new tools write the required data to three geodatabase tables that mirror the intermediate file structure. The .psr file is used by the WinHSPF system only if point sources are being considered, and this functionality is not supported in the ArcGIS system. The Geoprocessing tools assume that these three tables (wsdTable, rchTable, and ptfTable) are present in the geodatabase that contains the Land and River Segment GIS data.

Write to wsd Table in Database

The first of the Geoprocessing Tools developed for the ArcGIS HSPF Preprocessing system to create a new .uci file uses information from the Land Segment feature class and the HSPFLuseClass table to write the information for the .wsd text file. Figure 5.12 shows the interface for this tool. The two primary input parameters are the Land Segment feature class and the HSPF Landuse Classification table. Additional input parameters are available to choose the fields that contain pertinent information. These parameters will default to values from the ArcGIS HSPF Preprocessing geodatabase but can be changed by the user. The tool assumes a table named “wsdFile” is present in the geodatabase containing the HSPFLandSegment feature class.

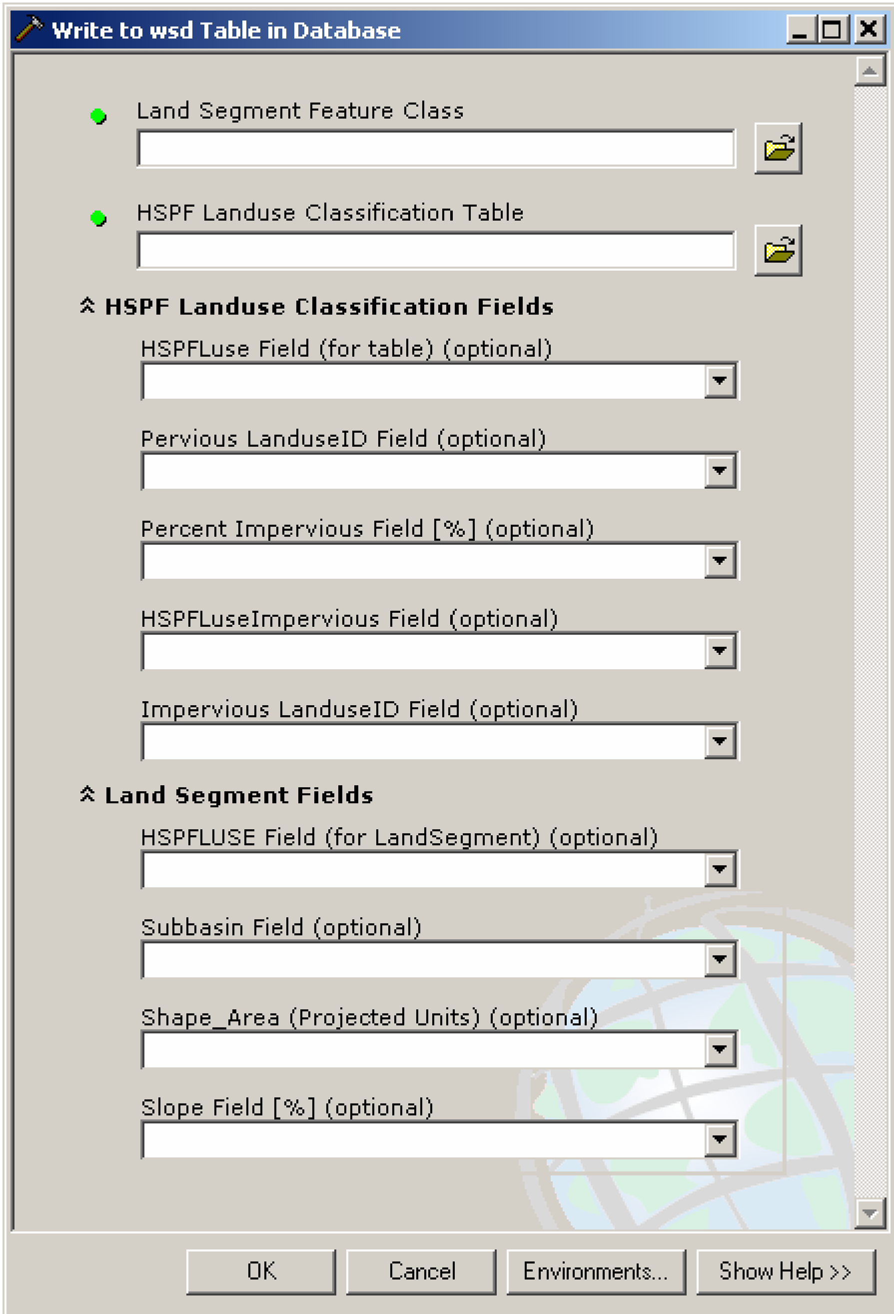


Figure 5.12 Custom ArcGIS Geoprocessing function: Write to wsd Table in Database.

The algorithm behind the tool loops through each feature in the HSPFLandSegment feature class and queries the HSPFLuseClass table to find the type and amount of impervious land associated with the HSPF landuse type. If there is no impervious area associated with a landuse type, the tool simply writes a single line to the wsdTable in the database with the area of the feature (in acres). If impervious area is associated with a landuse type, the tool writes two lines to the .wsdTable in the database, one defining an impervious Land Segment and the other defining a pervious Land Segment. The appropriate areas are calculated using percent impervious associated with the landuse type in the HSPFLuseClass table. Details of how attributes from the LandSegment and HSPFLuseClass table are used to write lines in the .wsd text file are presented in Appendix D.

Lines are added to the wsdTable in the order of the Subbasin field. In order to ensure a consistent and predictable Operation Numbering convention in the resulting .uci file, some additional lines are written to the .wsd text file. For the first subbasin that the tool encounters (the one with the lowest subbasin attribute), a line is added to the .wsd table for each type of landuse defined in the HSPFLuseClass table whether or not there is any of this type of land in the subbasin. Landuse types that are not present in this first subbasin are assigned a contributing area of zero, so they do not affect the actual simulation. By carefully allocating each type of land to be simulated in the HSPF model in the first subbasin, the third digit of the Operation Number will be consistent with that defined by the PLuseID attribute in the HSPFLuseClass table in the geodatabase.

Fields required for the tool default to values from the ArcGIS HSPF preprocessing database, but can be changed by the user if the necessary information is stored in fields with different filenames. The required fields for the Land Segment feature class include:

1) the HSPFLUSE Field, containing a string corresponding to the HSPFLUSE field in the HSPFLuseClass table
2) the Subbasin Field, containing a long integer corresponding to the River Segment that the land area contributes to
3) Shape_Area, contains the area of the feature in the projected units

4) the Slope field, contains the average subbasin slope or the slope of the individual land segment (in %)

Required fields for the LuseHSPFClass table include:

1) the HSPFLUSE Field, containing a string corresponding to the HSPFLUSE field in the Land Segment feature class
2) the Pervious LanduseID field, containing a long integer (1-9) used as the third digit for the Operation Numbering convention for Pervious Land Segment Operations. Should only use 1-9.
3) the Percent Impervious field, containing a long integer defining the percent of impervious land associated with the landuse type
4) the HSPFLUSEImpervious Field, containing a 20 or less character string defining which impervious land category the landuse type belongs to
5) the Impervious LanduseID field, containing a long integer (1-9) used as the third digit for the Operation Numbering convention for Impervious Land Segment Operations. This number should be the same for all records with the same HSPFLUSEImpervious string and be between 1 and 9.

Write to rch and ptf Tables in Database

The second of the Geoprocessing Tools for creating a new .uci file uses information from the River Segment feature class to write the information for the .rch and .ptf text files. Figure 5.13 shows the interface for this tool. The primary input is the DrainageLine feature class, whose attributes are used to define or calculate the information for both the .rch and .ptf text files. Additional input parameters are again available to choose the fields that contain pertinent information. These parameters will default to values from the ArcGIS HSPF Preprocessing geodatabase but can be changed by the user to accommodate data in a different format. The tool assumes two tables named “rchFile” “ptfFile” are present in the geodatabase containing the DrainageLine feature class.

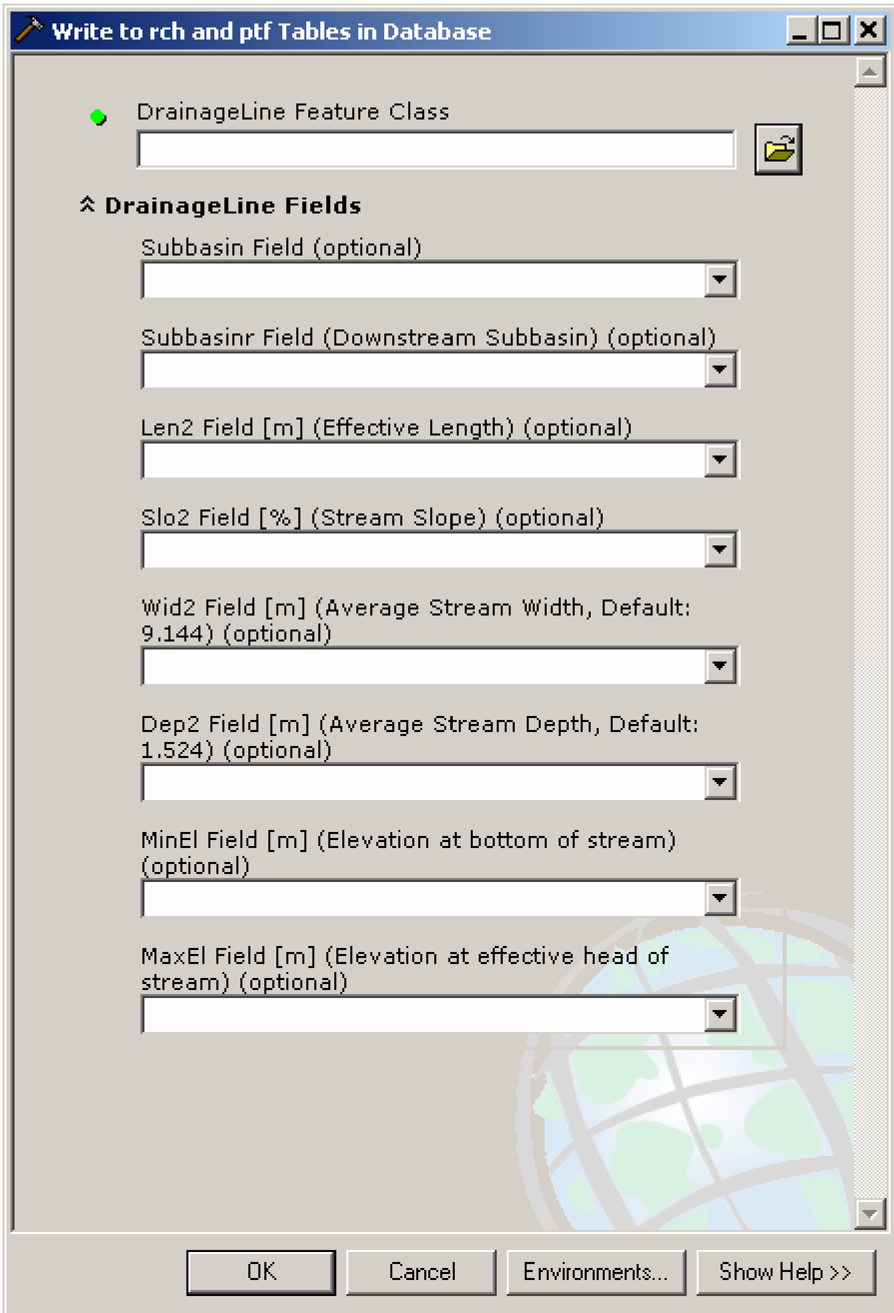


Figure 5.13 Custom ArcGIS Geoprocessing function: Write to rch and ptf Tables in Database.

The algorithm behind the tool uses the attributes of the DrainageLine feature class to populate the appropriate fields in the rch and ptf geodatabase tables. A single record is added to the rch and ptf geodatabase tables for each feature in the DrainageLine feature class. If the Width and Depth fields of the DrainageLine feature class do not contain non-zero values, defaults values of 30 feet for width and 5 feet for depth will be assigned. Details of how the other attributes from the DrainageLine feature class are used to write lines in the .rch and .ptf text files are presented in Appendix D.

Required fields for the DrainageLine feature class include:

1) the Subbasin Field, containing the long integer corresponding to the River Segment to which land area contributes
2) the Subbasinr Field, containing the long integer corresponding to the next downstream River Segment
3) the Effective Length Field, containing the effective length of the stream in projected units, which may or may not be the same as the Shape_Length attribute of the DrainageLine features
4) the Average Slope Field, containing the average slope of the river segment (in %)
5) the Average Width Field, containing the average width of the river segment (in meters). Will default to 30 feet if values are not specified in the DrainageLine feature class
6) the Average Depth Field, containing the average depth of the river segment (in meters). Will default to 5 feet if values are not specified in the DrainageLine feature class
7) the MinEl Field, containing the minimum elevation for the river segment (in meters)
8) the MaxEl Field, containing the elevation for the effective head of the river segment (in meters)

Write Text Files

The third and final tool used in the process of writing the intermediate text files writes the contents of the three geodatabase tables (wsd, rch, ptf) to a specified folder and assigns a project name. The interface to the tool is shown in Figure 5.14.

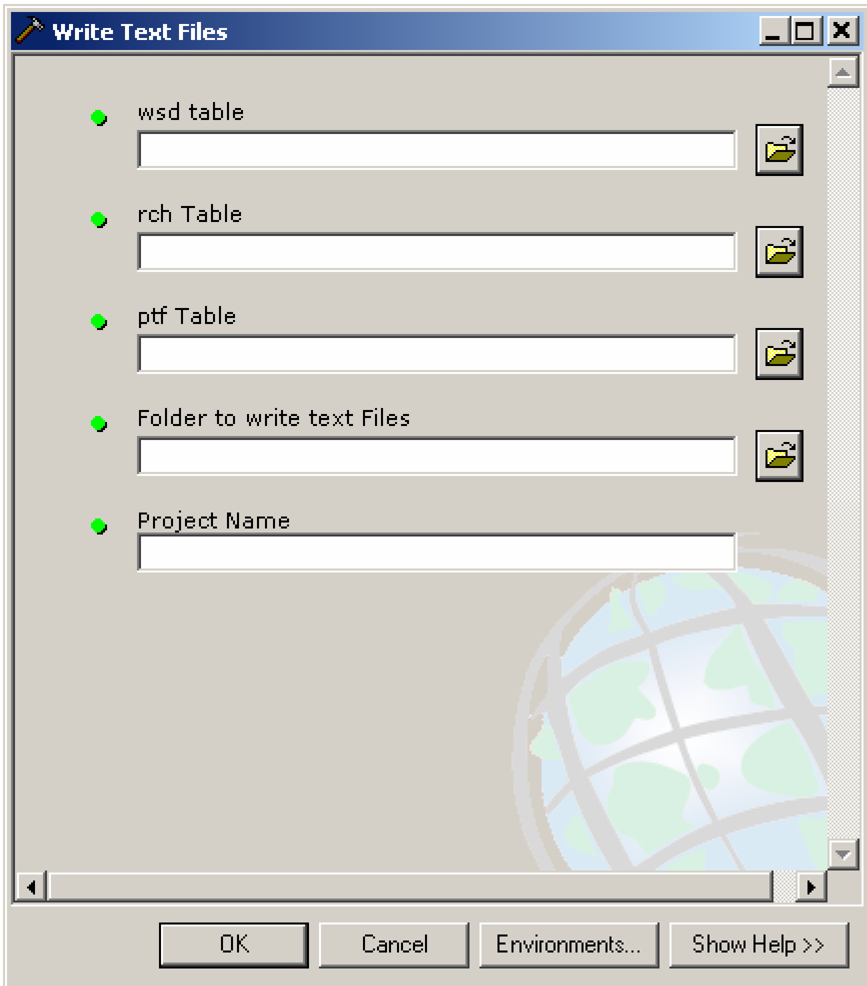


Figure 5.14 Custom ArcGIS Geoprocessing function: Write Text Files.

The tool takes three geodatabase tables as inputs as well as the workspace (folder) and project name for a new set of intermediate text files. The tool copies template intermediate files from a workspace specified as an additional input. These files are assumed to be named “newWsd.wsd,” “newRch.rch,” “newPtf.ptf,”and “newPsr.psr.”

Template files are copied to avoid the necessity of creating text files with the appropriate extensions.

Create new HSPF model using WinHSPF

The final process of creating a new .uci file involves using the WinHSPF program to create a new .uci file from the intermediate text files. From the WinHSPF interface, the "Create Project" function is chosen from the "File" menu. A "BASINS Watershed File" is selected and the tool automatically looks for other intermediate text files (.rch, ptf, and .psr) with the same filename and appropriate extensions. After .wdm datasets are specified as input to the model, a new .uci file is created with the same name as the intermediate text files. Figure 5.15 shows the WinHSPF program's "Create Project" tool interface.

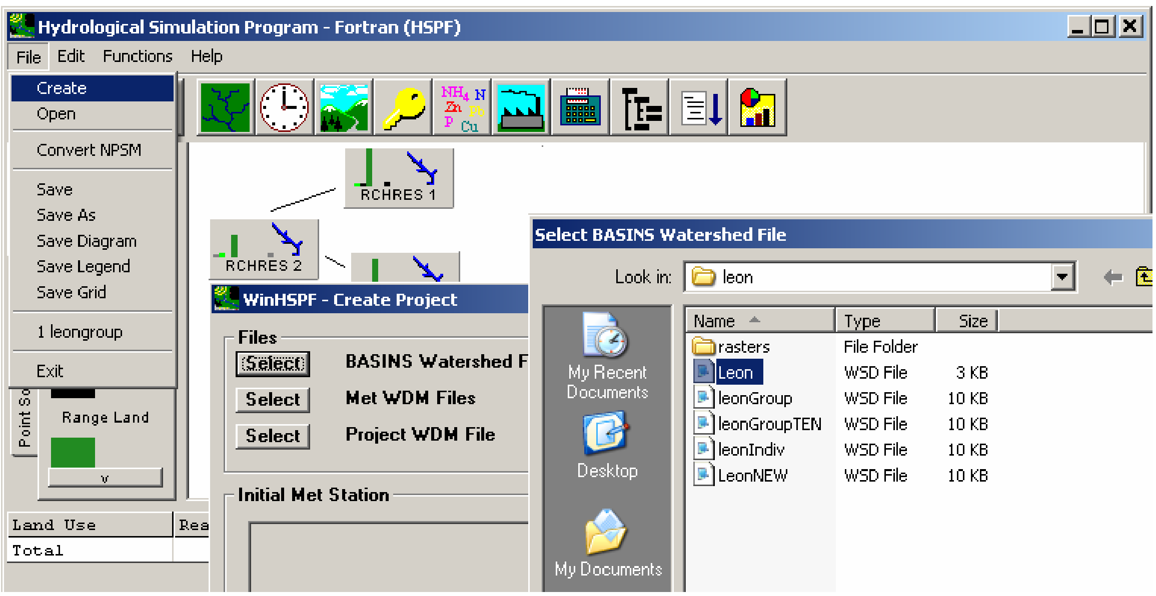


Figure 5.15 WinHSPF 'Create Project' function.

This final part of the ArcGIS HSPF Preprocessing methodology was not implemented in a Geoprocessing function because the functionality is already fully present in the WinHSPF program.

Copy Field: Table to Feature Class

The tool used to copy fields from tables to feature classes in some of the models presented above is not available in the standard ArcGIS tools, but was developed as a custom tool and included with the ArcGIS HSPF Preprocessing tools. The interface for this tool is shown in Figure 5.16. The tool takes two files as inputs, "Incoming Table (To Copy From)" and "Outgoing Table or Feature Class (To Copy To)."

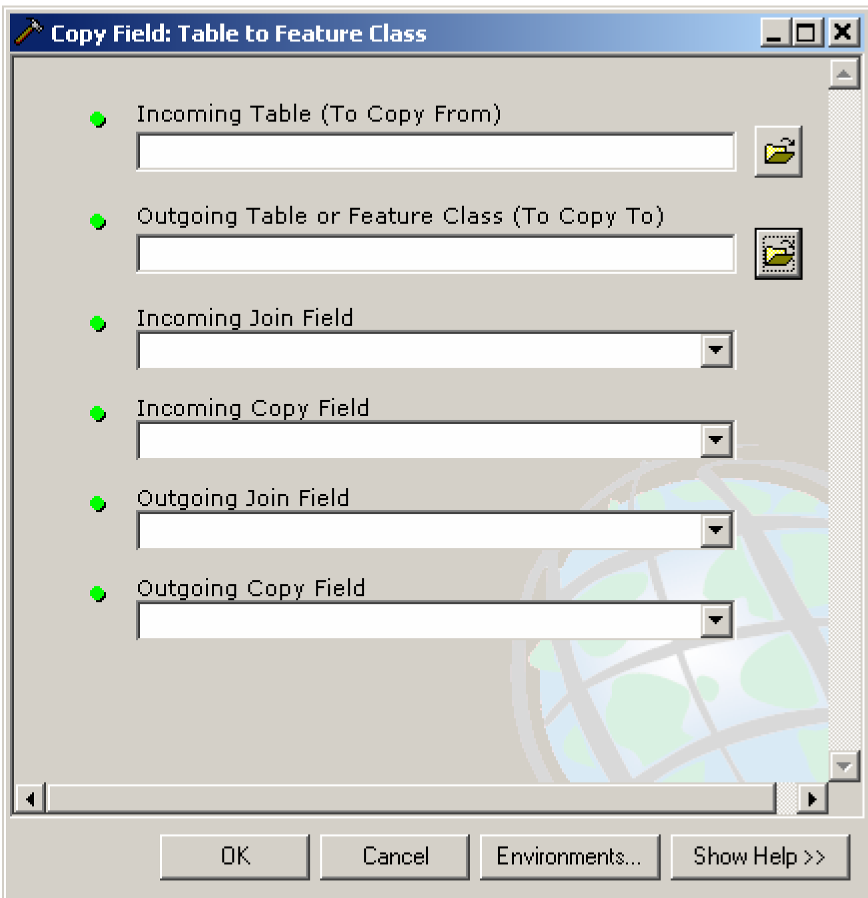


Figure 5.16 Custom ArcGIS Geoprocessing function: Copy Field: Table to Feature Class.

There must be a one-to-many relationship between records in the Incoming Table and the Outgoing Table or Feature Class, so that records from the Incoming Table can be used to copy a value to one or more records in the Outgoing Table or Feature Class. The "Incoming Join Field" is the field in the Incoming Table that has a single unique value for each value in the “Outgoing Join Field” in the Outgoing Table or Feature Class. The “Incoming Copy Field” contains the value that will be added to the Outgoing Table or Feature Class, and the “Outgoing Copy Field” is the field to which this new value will be written. The “Outgoing Copy Field” must be present in the Outgoing Table or Feature Class, so this tools often used in conjunction with the “Add Field” tool available in the standard DataManagement-Fields toolset in ArcToolbox.

This tool was developed to avoid using the standard ArcGIS “Join Based on an Attribute” tool which only operates on feature layers and not feature classes. This tool was developed specifically to map landuse categories from GIS data to those to be used in HSPF modeling with the LuseDefinition table (see Section 4.1.3), but is useful for many other processes as well.

5.4 GEOREFERENCING HSPF MODELS

The tools and models of the HSPF Preprocessing system presented thus far essentially mirror the BASINS HSPF Preprocessing system. Though they differ slightly in the data structures, the ArcGIS process follows the same methodology to define Land Segments using GIS landuse data, drainage area boundaries, and River Segments. In both systems the data is written to intermediate text files and used to build a new .uci file with WinHSPF.

An additional aspect of the ArcGIS HSPF Preprocessing system involves creating and maintaining a geospatial representation of Model Elements in GIS data. The HSPFCode (see Section 4.3.3) is a unique identifier on each Model Element both in the .uci file and in GIS data. The HSPFCode is assigned using a custom Geoprocessing function and subsequently used by the ArcGIS Timeseries preprocessing system to automatically update HSPF model files based on information in the geodatabase.

The GeoReferencing tool presented below assumes that the .uci file has been built using the WinHSPF 'Individual' option and that the ArcGIS HSPF Preprocessing methodology was used to develop the model.

5.4.1 Georeferencing Tool

A Geoprocessing tool was developed to assign this HSPFCode to Model Elements and the associated GIS data. The user interface to the tool is shown in Figure 5.17. The primary input parameters to the tool are:

a. the HSPFLandSegment feature class
b. the LuseHSPFClass table
c. the DrainageLine feature class
d. the .uci file (workspace and filename)

All the fields used by the tool will default to values from the HSPF Preprocessing geodatabase, but can be changed by the user if necessary.

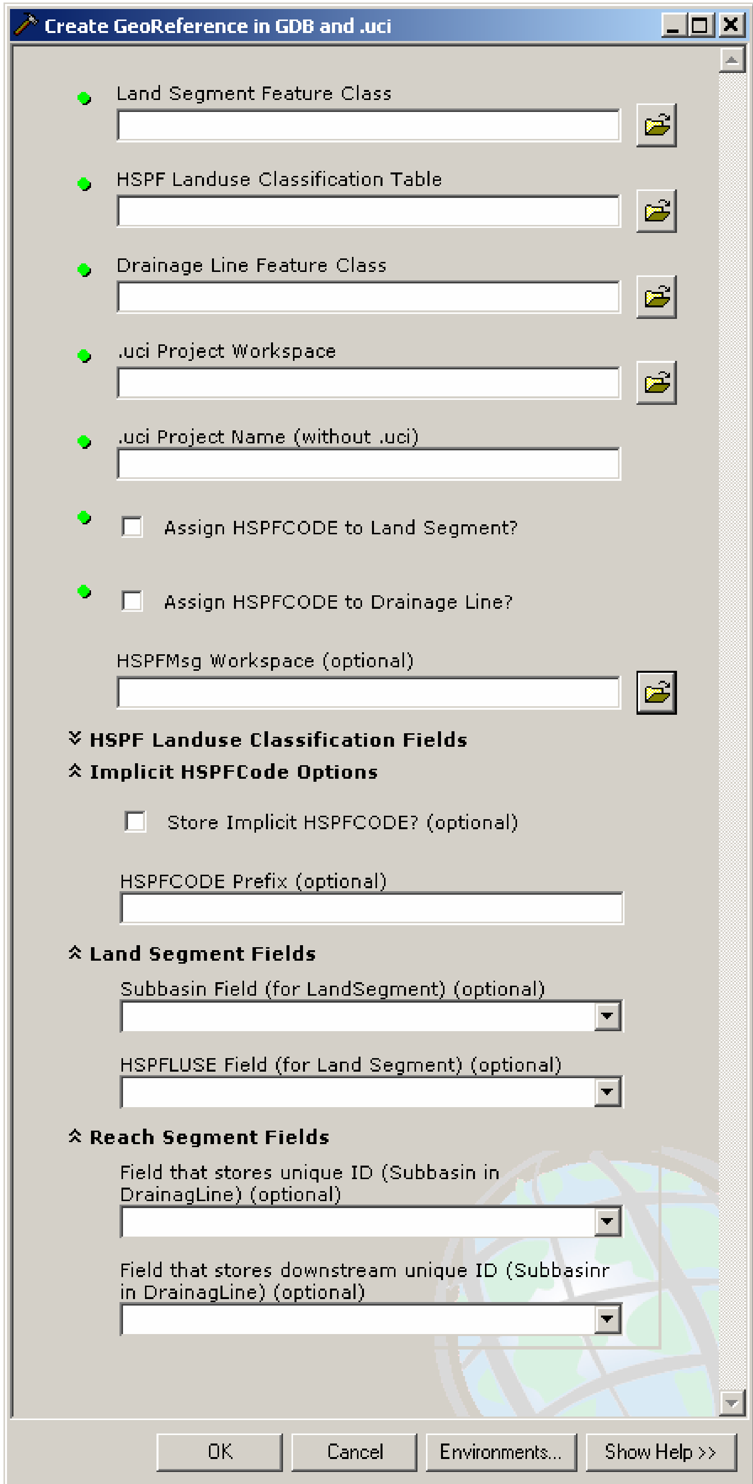


Figure 5.17 Custom ArcGIS Geoprocessing function: Create GeoReference in GDB and .uci.

The basic concept of the algorithm is to read the .uci file into the computer memory and loop through each model element (RCHRES, PERLND, and IMPLND). An HSPFCode is assigned to each Model Element in the .uci file based on the Operation Number and Operation Type. For each Model Element encountered in the .uci file, the tool queries the appropriate GIS data to find the corresponding GIS feature or features. The same HSPFCode is assigned to both the Model Element in the .uci file and its geospatial representation in the GIS data.

The tool is designed to work on .uci files created following the ArcGIS Preprocessing methodology and the “Individual” option of WinHSPF’s “Create Project” tool. For Land Segments, it assumes that the first two digits of the Operation Number correspond to the “Subbasin” attribute from the HSPFLandSegment feature class and the third digit corresponds to the type of land as defined by the HSPFLUSE attribute in the HSPFLandSegment feature class and the LuseHSPFClass table. For River Segments, the tool assumes the Operation Number corresponds to the Subbasin attribute from the DrainageLine feature class.

It should be noted that this tool only works properly if 10 or more River Segments are present in the model. This limitation is due to differences in the conventions adopted by WinHSPF for numbering Operations when creating a new model with 9 or fewer River Segments. When WinHSPF is used to create a new HSPF model using the ‘Individual’ option , the final digit in the Operation Number is used to distinguish the type of landuse simulated by the Operation, leaving the first two digits of the Operation Number for up to 99 River Segments. If 10 or more River Segments areas are present, the first two digits will be assigned two digits corresponding to the river segment to which they drain. If fewer than 10 uniform rain areas are present, the first two digits of Land Segment Operation Numbers are assigned “10x” thru “90x.” The tool to Create GeoReference in GDB and .uci assumes that the first two digits of the Operation number correspond directly to the River Segment numbers and will not work properly for models with 9 or fewer River Segments.

Because some features in the HSPFLandSegment feature class represent both pervious and impervious land, two attributes are used to store HSPFCodes on the LandSegment feature class. The "HSPFCODE" attribute stores the HSPFCode corresponding to the Pervious Land Segment, and the "HSPFCODEI" attribute stores the HSPFCode corresponding to the Impervious Land Segment. If these attributes do not exist when the tool is run, they will be added. In addition to the HSPFCode attributes, the tool also assigns attributes related to viewing HSPF binary output in the GenScn program. These attributes, "GnSnLocP" and "GnSnLocI" for the Land Segments and "GnSnLoc" and "GnSnDnID" for the River Segments correspond to the location attribute of the HSPF Binary Output files. Binary output is sometimes used to view output from HSPF in the GenScn program.

In addition to the primary input parameters mentioned above, several other inputs are used by the tool. The HSPFMsg.mdb database is used by the WinHSPF libraries to aid in reading .uci files into the program's memory. Its location is defaulted to "C:\BASINS\bin" unless otherwise specified. The database should have been created when the GenScn/WinHSPF/WDMUtil programs were installed in the "…\BASINS\bin" directory.

Options are available to assign HSPFCode and GenScn locations to only the DrainageLine feature class, only the LandSegment feature class or both. Another option allows the user to store the HSPFCode implicitly and is used by a specific application of the tools in the San Antonio area.

For the tool to read the .uci file into memory and work properly, the location of all .wdm files specified in the FILES block of the .uci file must contain the entire pathname for the file. Relative path names are supported for WinHSPF and WinHSPFLt, but are not sufficient for the new tool. Explicit, complete pathnames must be present when the tool is initially used even though relative pathnames may be written back to the .uci file upon completion of the tool.

5.4.2 HSPFCode: Details and Structure

The tool described above uses the Operation Number as a basis for assigning an HSPFCode to the Model Elements in the .uci file and the GIS data. The Operation Number provides a convenient starting point for the HSPFCode because it is often used to implicitly store geospatial information in the HSPF .uci file. The HSPFCode concept is not inherently related to the Operation Number, however, and it could be any identifier (numeric or alphanumeric) that is present both on Model Elements in the .uci file and in GIS data.

The tool presented above assigns an HSPFCode to each model element based on the Operation Number. Tools developed for the ArcGIS Timeseries Preprocessing system (presented in the next Section), however, are designed to use the HSPFCode and not the Operation Number to transfer information between GIS data and the .uci file. These tools (from the Timeseries Preprocessing system) assume the HSPFCode for each Model Element is stored in the LSID (for PERLND and IMPLND) and RCHID (for RCHRES) field of the GEN-INFO block in the .uci file and are not dependent upon the Operation Numbering convention of the HSPF .uci file.

Two options for assigning the HSPFCode are available in the tool described above; one stores the HSPFCode explicitly, and the other, implicitly. The HSPFCode is stored explicitly if the box for “Store Implicit HSPFCode” is left un-checked. With the explicit option (box unchecked), the same HSPFCode is placed in both the .uci file and the geodatabase. The same character string is used to uniquely identify model elements in both the .uci file and the geodatabase. This character string is assigned as the first letter of the Operation Type concatenated with the three-digit numeric Operation number including preceding zeroes. Examples include “P101,” for PERLND 101, “I011” for

IMPLND 11, and "R010" for RCHRES 10. In the .uci file, this HSPFCode is written to the first few characters of the LSID/RCHID field of the GEN-INFO block of text lines. For Land Segment features in the geodatabase, the HSPFCode is written to a string attribute called "HSPFCODE" for PERLND Model Elements and to "HSPFCODEI" for IMPLND Model Elements. For the geospatial representation of RCHRES Model Elements, the HSPFCode is written to an attribute called "HSPFCODE."

If the HSPFCode is stored implicitly, a slightly different but related character string is used to identify Model Elements in the geodatabase and .uci file. With the implicit option (box checked) the entire HSPFCode is not stored in the LSID/RCHID fields of the .uci file. Instead, only the Operation Number is stored in the LSID/RCHID field. A one or two character alpha prefix is assumed depending on whether the element is a PERLND, IMPLND, or RCHRES Operation; "P," "I," or "QR," respectively. In addition, the first five characters of the RUN-INFO line near the top of the .uci file is assigned another set of characters that are implicitly part of the HPSFCode. The HSPFCode written to the GDB includes all three parts of the HSPFCode stored in the .uci file, the assumed alpha prefix, the RUN-INFO prefix, and the Operation Number. Table 5.1 summarizes the differences between storing the HSPFCode explicitly and implicitly.

Table 5.1 Explicit and Implicit HSPFCode options.

		PERLND	**IMPLND**	**RCHRES**
Explicit	***Geodatabase***	P101	I011	R010
	UCI-LSID/RCHID Field	P101	I011	R010
	UCI-RUN-INFO		--	
	UCI-Implicit Alpha Prefix	--	–	–
Implicit	*Geodatabase*	P10101101	I10101011	QR10101010
	UCI-LSID/RCHID Field	101	011	010
	UCI-RUN-INFO		10101	
	UCI-Implicit Alpha Prefix	P	I	QR

Note: For 5 digit Run-Info prefix "10101"

The explicit option for storing the HSPFCode provides a much more intuitive, straightforward way of linking the HSPF Model Elements and their GIS representation. Tools developed for the ArcGIS Timeseries Preprocessing system (presented in the next section) assume an explicit link between the elements of an HSPF model and their geospatial representation in the geodatabase.

The tools presented in Section 5.5.2 assume that a unique identifier (HSPFCode) for each Model Element resides in the LSID/RCHID field of the GEN-INFO section of the PERLND, IMPLND, and RCHRES Operation blocks. The same set of characters is assumed to uniquely identify the element in both the geodatabase and the .uci file. The HSPFCode is assumed to start at the first (leftmost) character of the LSID/RCHID field and go to the first "space" character (" " – ASCII 32). This means that the HSPFCode could be any set of up to 20 alphanumeric characters excluding spaces. HSPFCodes with spaces are not supported by the ArcGIS Timeseries Preprocessing tools.

5.5 ArcGIS Timeseries Preprocessing Implementation

The ArcGIS Timeseries Preprocessing system includes:

1) An extended geodatabase design to store and manage data during the development of input time series for an HSPF model
2) Custom tools to write time series to a .wdm file and update a .uci file to read from this new .wdm file.

The implementation depends upon time series data being stored in Arc Hydro format in the geodatabase with appropriate entries in the TSType and Timeseries tables. In addition, the HSPF model .uci file and corresponding GIS data must have HSPFCodes assigned as described in Sections 4.3.3 and 5.4.2.

The ArcGIS Timeseries Preprocessing methodology has been implemented to prepare NEXRAD precipitation data to drive HSPF models. NEXRAD data is typically stored in a gridded file format, with a single file storing data at regular spatial intervals for a large spatial domain, but only a single point in time. Wdm files on the other hand, require timeseries to be organized in a format in which individual datasets describe a temporal series of values at a single spatial location. The Arc Hydro timeseries format is different from either of these two formats. Because it is structured around relational database concepts, Arc Hydro timeseries is capable of organizing data into either spatial or temporal datasets depending on the type of query used to retrieve data. The tasks used to extract NEXRAD data from its native format to Arc Hydro format are described in Appendix B.

5.5.1 Geodatabase Design

The extended geodatabase design to support time series preprocessing for HSPF adds one feature class and two tables to the design presented in Section 5.3.1. Figure 5.18 shows the extended geodatabase design with the "GISMetSegment" feature class, TSType table and Timeseries tables.

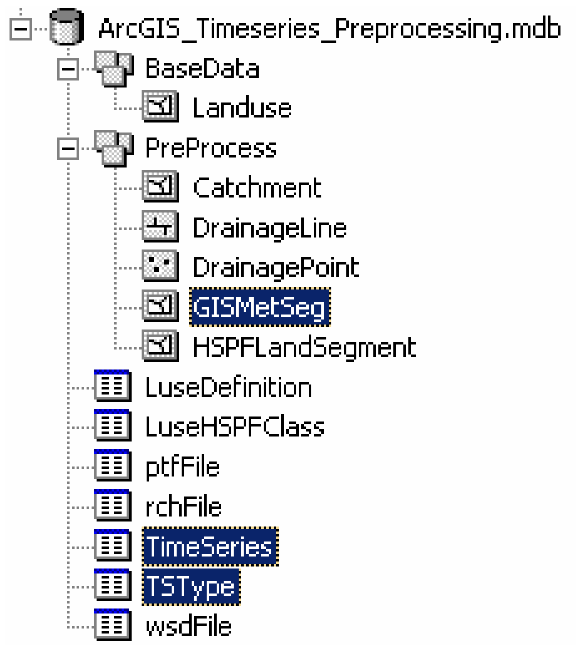


Figure 5.18 ArcGIS Timeseries Preprocessing geodatabase structure.

The fields in the Timeseries and TSType tables are the same as the standard Arc Hydro time series tables and are reviewed in Section 2.5.3, and in Maidment (2002). The Timeseries table contains all the time series values as well as information about which spatial feature the value is associated with. The TSType table contains other metadata information such as the type of data, units, and origin of the data.

The central data structure in the implementation of the ArcGIS Timeseries Preprocessing methodology is the GIS MetSegment feature class. GIS MetSegments are a polygon feature class representing areas of land over which areally uniform precipitation, evaporation and other atmospheric forcing data can be assumed. The ArcGIS HSPF Preprocessing methodology uses Arc Hydro timeseries data attached to this GIS implementation of the WinHSPF MetSegment concept to organize and prepare precipitation data for HSPF modeling. This discussion will focus on precipitation, but the concepts are general enough to be used with any of the eight time series associated with MetSegments.

Figure 5.19 shows the GIS data used in the ArcGIS Timeseries preprocessing system. GIS MetSegments are stored in the feature class called “HSPFMetSeg.” The “HSPFLandSegment” feature class contains a geospatial representation of each Model Element in an HSPF model, and the “Timeseries” table contains precipitation data in the Arc Hydro format. Each GIS MetSegment feature uses three types of attributes to communitcate with three different file types:

1) Arc Hydro time series with HydroID = TimeSeries.FeatureID)
2) .wdm files with attributes specifying .wdm datasets (i.e. PREC = dataset to store precipitation time series in .wdm file)
3) .uci files with the MetSegID = HSPFLandSegment.MetSegID)

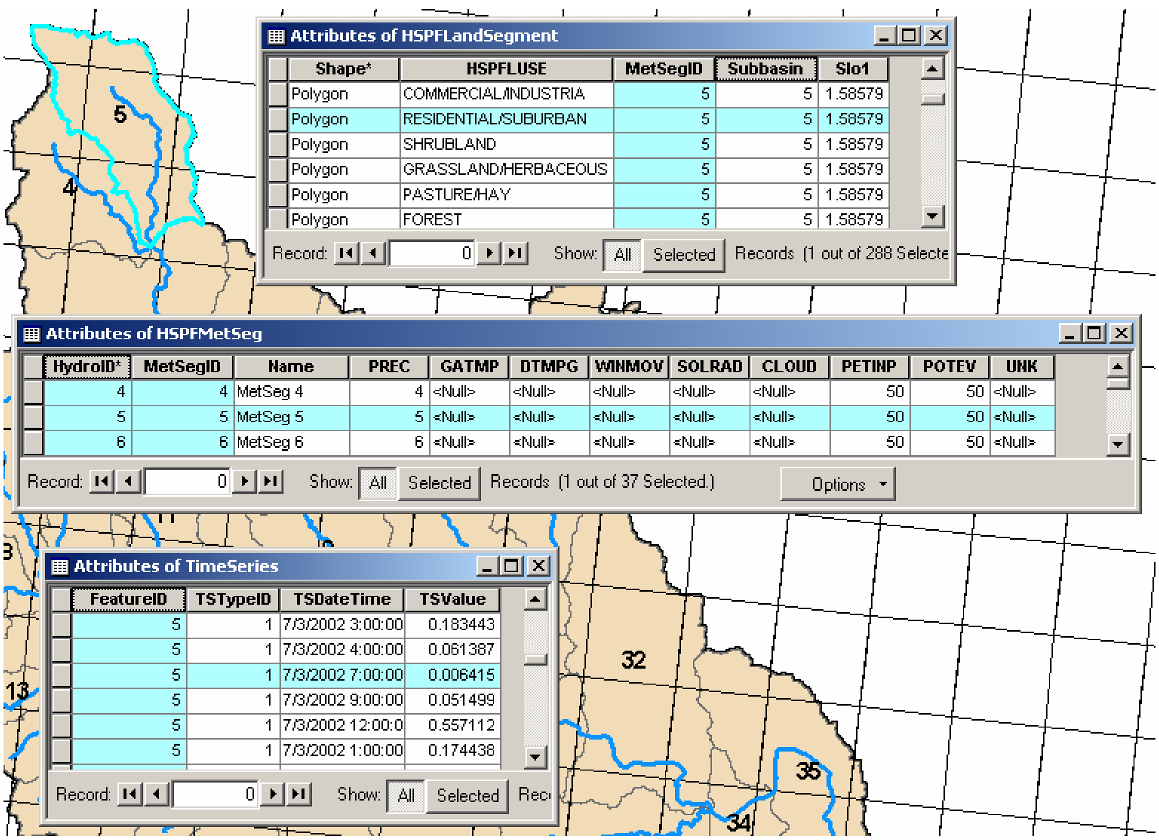


Figure 5.19 GIS MetSegment: Center of ArcGIS Timeseries Preprocessing methodology

Once the appropriate time series values are located and extracted from the geodatabase, they are written to the datasets specified in the appropriate attributes. The GIS MetSegment highlighted in Figure 5.19 (MetSegID = 5) is associated with the highlighted time series record where "MetSegID.HydroID = TimeSeries.FeatureID". Precipitation data associated with this GIS MetSegment feature will be written to .wdm data set #5.

Figure 5.19 shows that three GIS MetSegment features all specify data set #50 for both PETINP (potential evaporation over land segments) and POTEV (potential evaporation over river segments). If a single timeseries will be used to simulate the area of more than one GIS MetSegment feature, the timeseries will only be written once.

Figure 5.20 provides a summary of the attributes for the "GISMetSegment" feature class in the ArcGIS Timeseries Preprocessing geodatabase.

- **GISMetSeg feature class**
  - *Arc Hydro Attribute*
    - HydroID
  - *MetSegment ID Attributes*
    - MetSegID
    - Name
  - *MetSegment DSN Attributes*
    - PREC
    - GATMP
    - DTMPG
    - WINMOV
    - SOLRAD
    - CLOUD
    - PETINP
    - POTEV
    - UNK
  - *ArcGIS Attributes*
    - Shape_Length
    - Shape_Area

Figure 5.20 ArcGIS Timeseries Preprocessing feature class attributes: GIS MetSegment.

"HydroID" is used to connect the GIS MetSegment feature with the associated time series in the Timeseries table. Two attributes, "MetSegID" and "Name," define the

MetSegment; and nine additional attributes are used to specify the data set number to which corresponding time series is to be written in a .wdm file. Each DSN (data set number) attribute corresponds to one of the time series associated with MetSegments in the WinHSPF program and tools.

5.5.2 ArcToolbox Implementation

Two new Geoprocessing tools were developed as a part of the ArcGIS Timeseries Preprocessing system. The first tool extracts time series associated with the GISMetSegments from the geodatabase and writes it to the appropriate .wdm data sets based on the DSN attributes of the GISMetSegment feature class. The interface for this tool is shown in Figure 5.21.

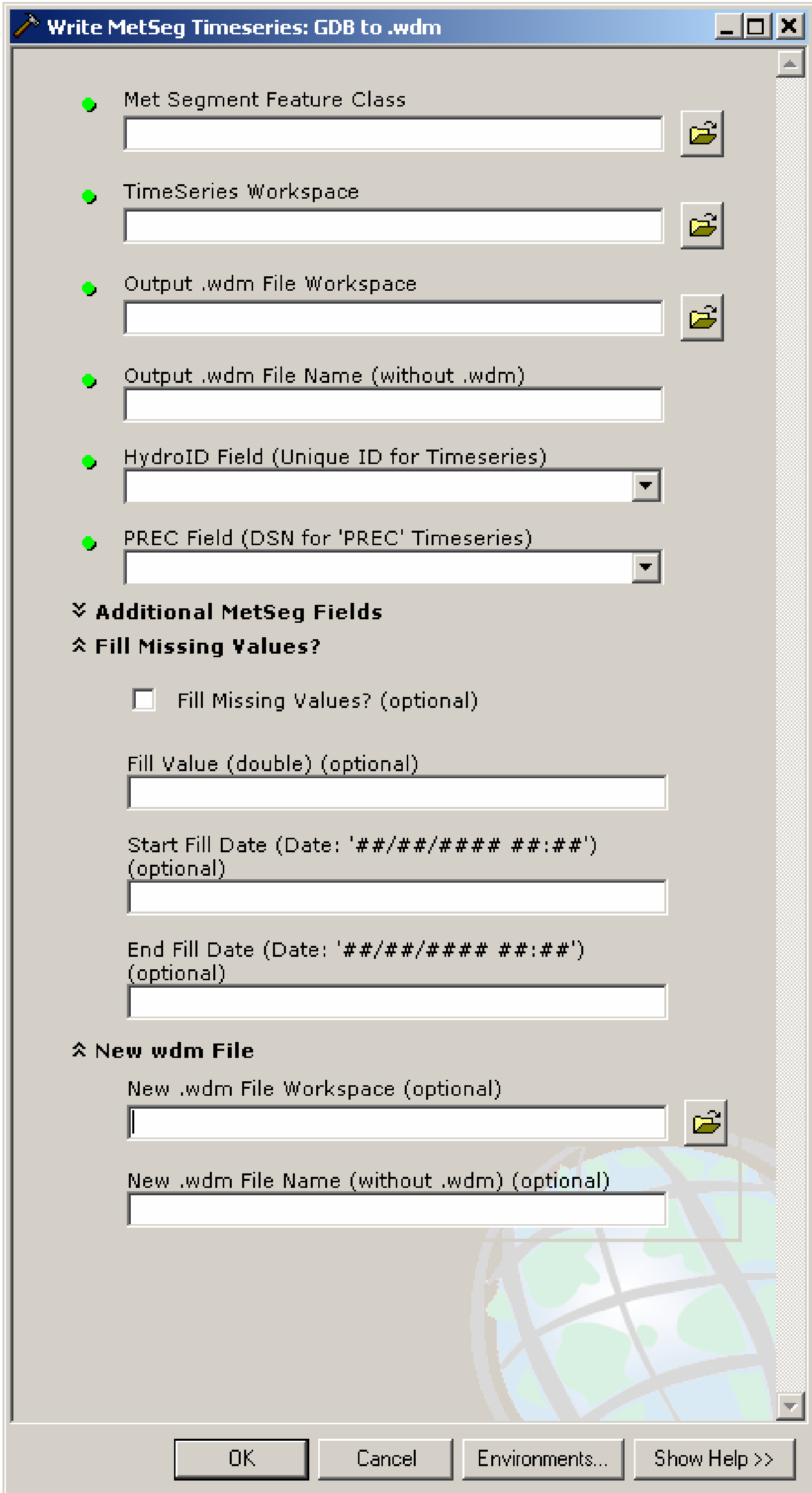


Figure 5.21 Custom ArcGIS Geoprocessing function: Write MetSeg Timeseries: GDB to .wdm.

The tool takes as the primary inputs the GISMetSegment feature class, the geodatabase containing the time series data, and an output .wdm file (workspace and filename). The "HydroID" field, used to link the geospatial features with time series data, and the field storing the datasets to write precipitation data are required at the minimum. Additional fields for other MetSeg time series DSN's can be assigned to read time series other than precipitation.

The tool first reads all the time series metadata (TSType table) into the program's memory and then loops through each feature in the GISMetSeg feature class and looks for non-zero values in the DSN attributes. If a non-zero value is encountered, the metadata for the Arc Hydro time series in the program memory is searched to find the appropriate values from the geodatabase. If the time series data is found, the data is assigned the appropriate data set number and written to the .wdm file. If the same DSN is assigned to more than one MetSegment feature, the data will only be written once.

To ensure that the tool works properly, the "Variable" field in the Arc Hydro TSType table should be assigned the appropriate MetSeg type, "PREC," "GATMP," "DTMPG," etc. For the two evaporation time series "POTEV" and "PETINP," "EVAP" can be used in the Arc Hydro TSType "Variable" field to accommodate evaporation for both Reach and Land Segments.

The tool does not create an entirely new .wdm file, but simply makes a copy of an existing, blank .wdm file in the location specified by the "Output .wdm File Workspace" and "Output .wdm File Name" parameters. The location of this blank .wdm file will default to "C:\ArcHydro2HSPF\bin" and "new" unless otherwise specified. Another option available in the tool allows for missing values to be filled. If the check box next to "Fill Missing Values?" is checked, the value specified in the "Fill Value" parameter will be included at every timestep between the "Start Fill Date" and the "End Fill Date" that did not contain a value in the Arc Hydro time series geodatabase. If the specified "Start Fill Date" or "End Fill Date" fall between the first and last dates from the Arc Hydro time series, the original start and end date will be retained. The tool does not support extracting only a portion of time series from a database as this can be easily done with an SQL query or in Microsoft Access.

The second tool developed for the ArcGIS Timeseries Preprocessing system updates the HSPF.uci file to assign the correct forcing time series to each Model Element. The interface for this tool is shown in Figures 5.22 and 5.23. Primary inputs are the LandSegment, Reach Segment, and MetSegment feature classes, the HSPF Model .uci file (workspace and filename), and the input .wdm file (workspace and filename). Three DSN attribute fields are required to run the model, PREC, PETINP, and POTEV. These correspond to the minimum precipitation and evaporation time series required to run an HSPF hydrology simulation. HSPFCode and MetSeg fields will default to values from the ArcGIS Timeseries Preprocessing geodatabase, but can be changed if necessary.

The algorithm behind the tool assumes that the HSPF .uci file and the GIS data used to develop it have been Georeferenced as described in Section 5.4.2. The tool reads the .uci file into the program memory and loops through all HSPF Model Elements in the .uci file, PERLND, IMPLND, and RCHRES. For each Model Element, the tool queries the corresponding GIS data to find which MetSegment is associated with the Model Element, as defined by the MetSegID attribute. Each GIS representation of a Model Element must have a "MetSegID" attribute that defines which MetSegment feature it is associated with.

As mentioned in Section 5.4.2, the tool assumes that a unique character string is stored in the LSID/RCHID field of each Model Element in the .uci file and that the geospatial representation of this model element also contains the same character string in an attribute of the associated GIS data. If multiple GIS features exist for a single model element, the tool will use the MetSegment associated with the first feature it encounters. If multiple features in the geodatabase represent the same Model Element, they should all be associated with the same MetSegment to ensure that the proper MetSegments are assigned.

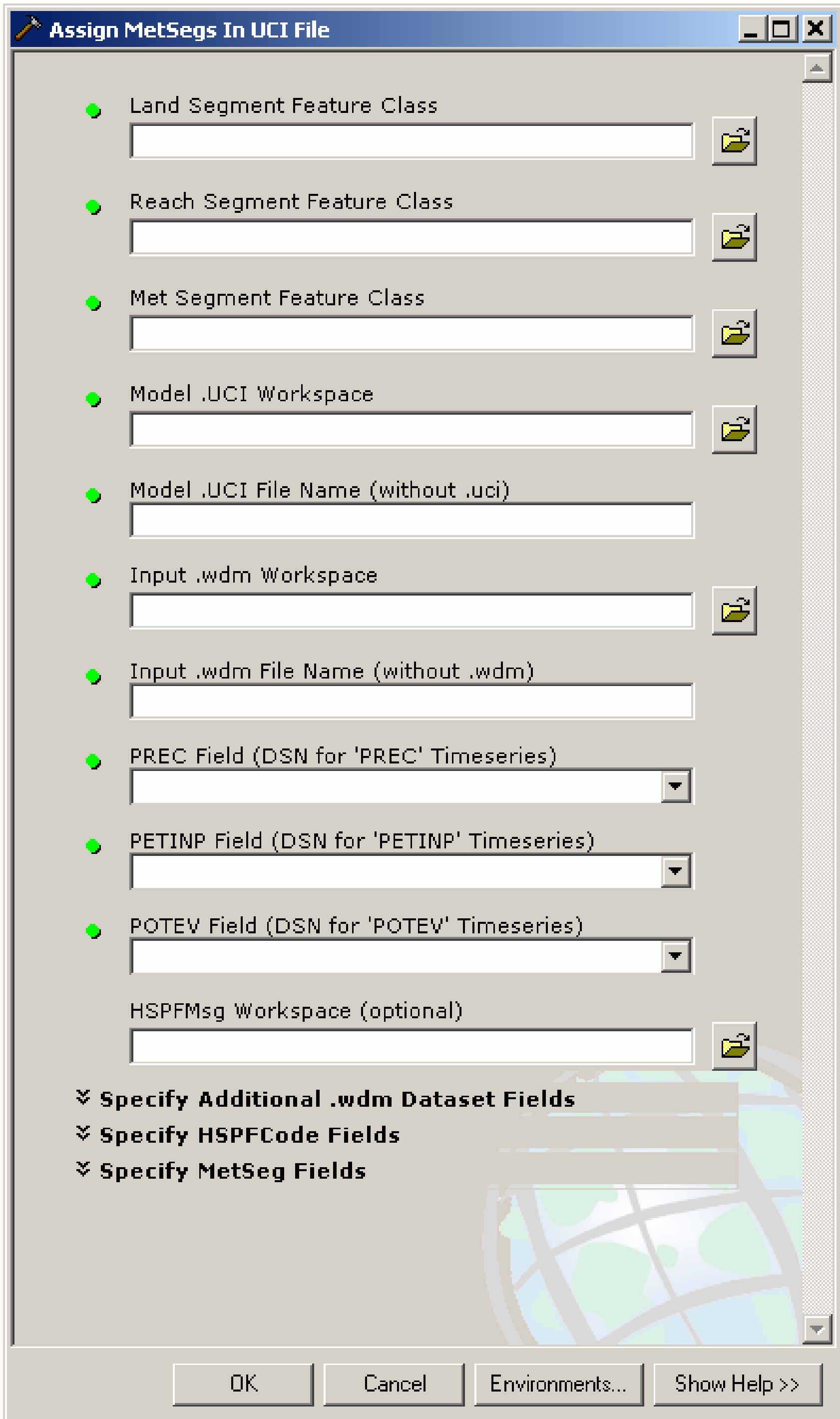


Figure 5.22 Custom ArcGIS Geoprocessing function: Assign MetSegs in .uci file (1).

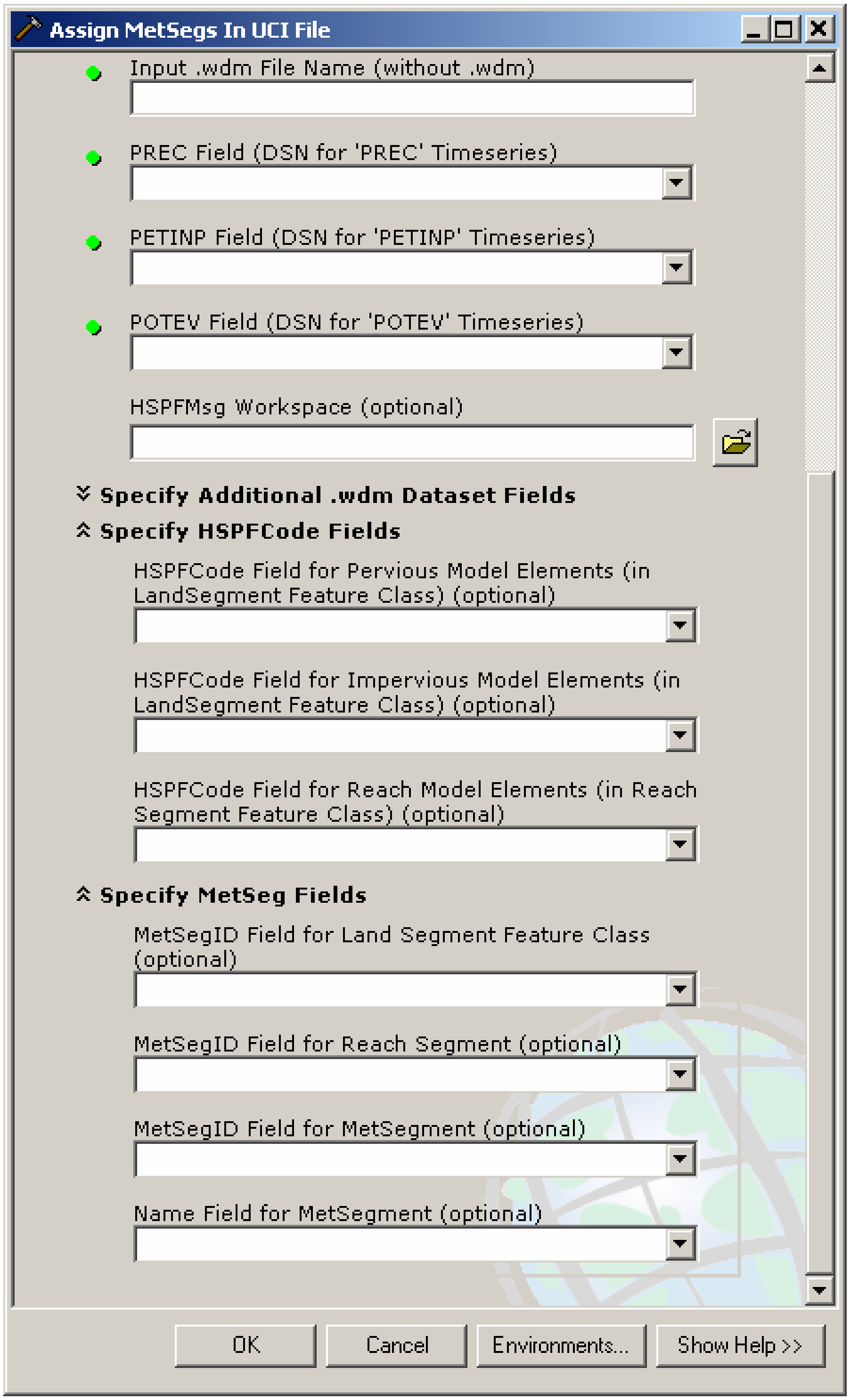


Figure 5.23 Custom ArcGIS Geoprocessing function: Assign MetSegs in .uci file (2).

5.6 SUMMARY

A structured geodatabase design and a set of existing and custom ArcGIS tools are implemented to prepare GIS data for use in HSPF model development. Beginning with Arc Hydro data for streams and drainage areas, standard ArcGIS tools perform the spatial analysis necessary to define the areas of land to be simulated by HSPF. Custom ArcGIS Geoprocessing functions extract the necessary information from GIS data to text files and WinHSPF is used to create a new HSPF .uci file.

A geospatial representation of HSPF Model Elements is maintained in the geodatabase during model development. This geospatial representation is used to transfer information from GIS data to HSPF model files. Precipitation timeseries stored in GIS data are written to .wdm files and a .uci file is updated accordingly.

Chapter 6 Results and Discussion

The primary result of this work is a standard methodology designed to facilitate the development of HSPF models in the ArcGIS environment. For HSPF modelers who wish to use ArcGIS instead of the ArcView 3.x BASINS system, the ArcGIS HSPF Preprocessing methodology provides a structured system for developing HSPF models. These tools are described elsewhere in this document and they will not be presented in detail here.

A secondary result of this work has been the development of HSPF models for several study areas near San Antonio, Texas. While these models could have been developed using the ArcView BASINS 3.x system or manually using ArcGIS data, the efficiency with which a standard set of well-structured HSPF models were developed with the aforementioned set of tools is a significant result of this work. The first two sections of this chapter are devoted to describing the application of the ArcGIS HSPF and Timeseries Preprocessing methodologies to several study areas. The final sections of this chapter discuss limitations encountered during the tool development and application and provide context for this work in the larger GIS and HSPF modeling community.

6.1 Results of ArcGIS HSPF Preprocessing Applications

Two separate projects have required the development of HSPF models at the Center for Research and Water Resources (CRWR) over the past year. The Bexar Regional Watershed Management Coalition's (BRWMC) Watershed Management Plan includes the use of HSPF models for simulating water quality in the San Antonio River and its tributaries. The project involves developing a standard, consistent set of HSPF models that will be available to interested parties for water quality simulations.

A second project underway at CRWR involves the development of Total Maximum Daily Loads (TMDL) for two impaired rivers near San Antonio. This project is still underway, but will eventually use HSPF to simulate dissolved oxygen and bacteria processes on the watersheds contributing to the impaired segments.

The ArcGIS HSPF Preprocessing system (presented in Chapters 4 and 5) has been used to develop HSPF models for both of these projects. For the BRWMC project, models have been developed for three watersheds: Cibolo Creek, Leon Creek, and Rosillo Creek. HSPF models are under development for TMDL's in the Sandies and Elm and Atascosa watersheds. Figures 6.1 shows the locations of the study areas.

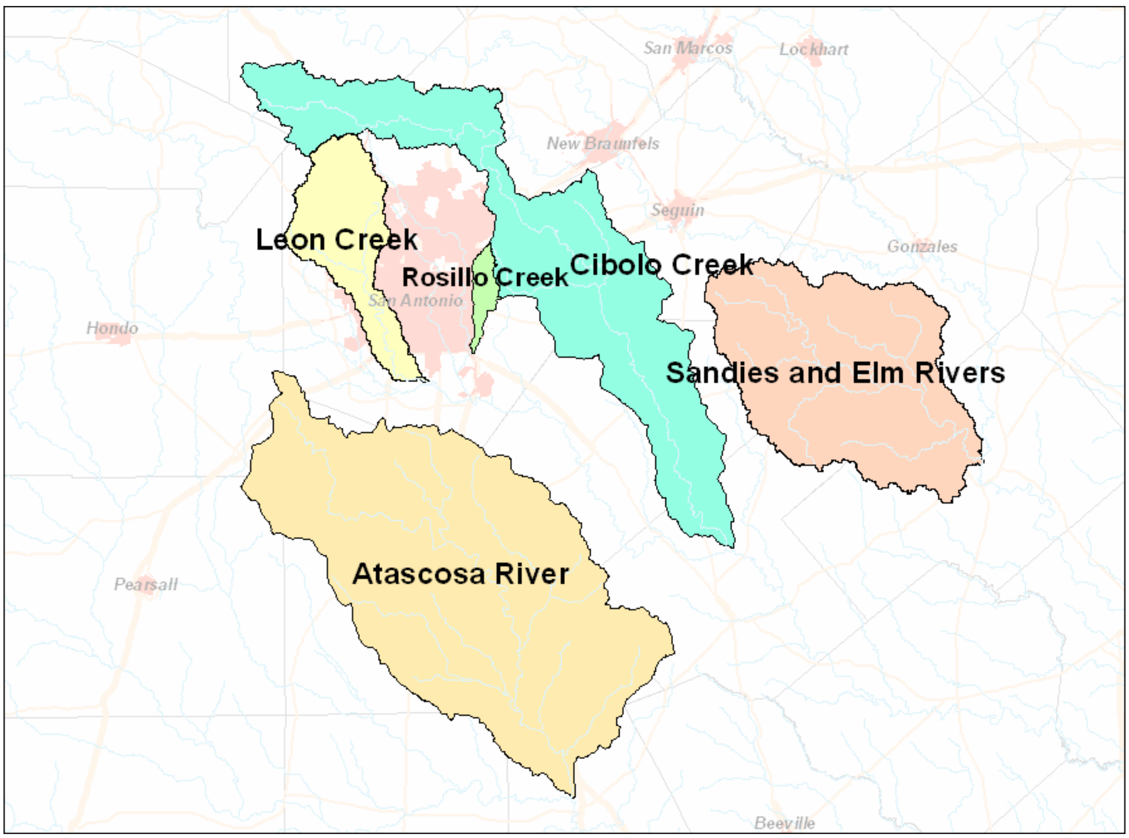


Figure 6.1 Five study areas for application of ArcGIS HSPF Preprocessing tools.

For each study area, the ArcGIS HSPF Preprocessing methodology was applied as outlined in Chapter 5.

1) Begin with Arc Hydro DrainageLine and Catchment data
2) Calculate physically based parameters for HSPF model creation
3) Define Land Segments to be simulated by HSPF
4) Extract necessary information to intermediate text files
5) Use WinHSPF to build a new HSPF model

In each application, digital elevation models were used to delineate drainage areas for the river network using the Arc Hydro tools and the results were loaded into the ArcGIS HSPF Preprocessing geodatabase. Figure 6.2 shows the locations of the Arc Hydro data used to begin HSPF model development.

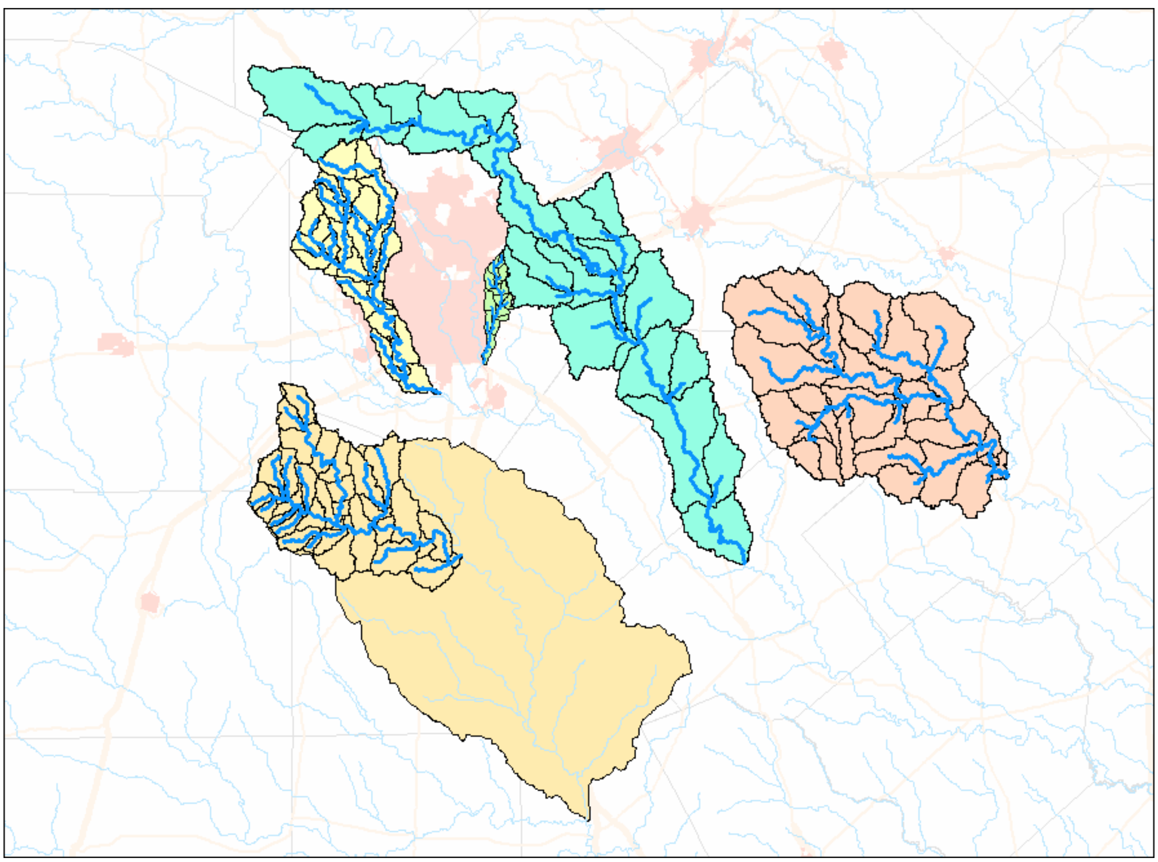


Figure 6.2 Arc Hydro GIS data used to begin HSPF Model Development.

When the Arc Hydro data is first loaded into in the ArcGIS HSPF Preprocessing geodatabase, only the attributes from the Arc Hydro system are present as shown in Figure 6.3. Tools from the ArcGIS HSPF Preprocessing methodology, presented in Chapters 4 and 5, are used to calculate the attributes required for HSPF model creation. The model development process is presented in the following figures only for Leon Creek; however, the same process was followed for each of the five study areas.

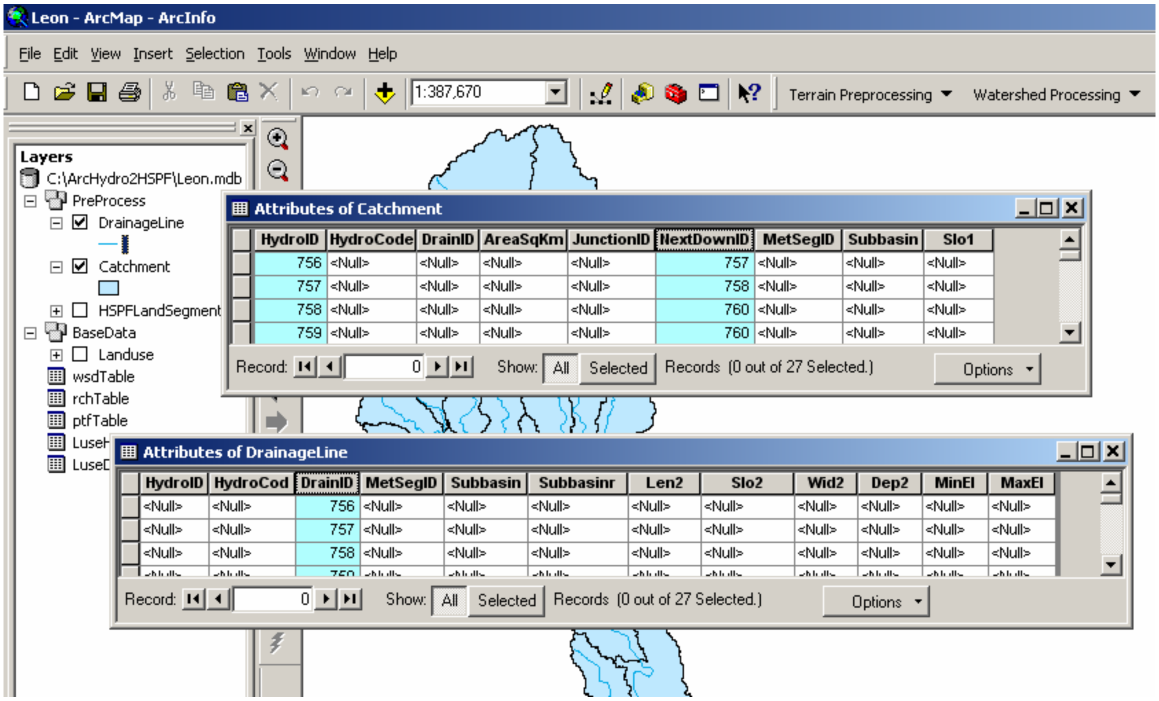


Figure 6.3 Arc Hydro feature classes before calculating attributes for HSPF model development.

Figure 6.4 shows the DrainageLine and Catchment features after the application of the ArcGIS HSPF Preprocessing tools. Catchments and DrainageLines are renumbered and network information is transferred from the HydroID-NextDownID attributes to the Subbasin-Subbasinr attributes. Slopes and Lengths are calculated using standard ArcGIS tools and placed in the appropriate attributes.

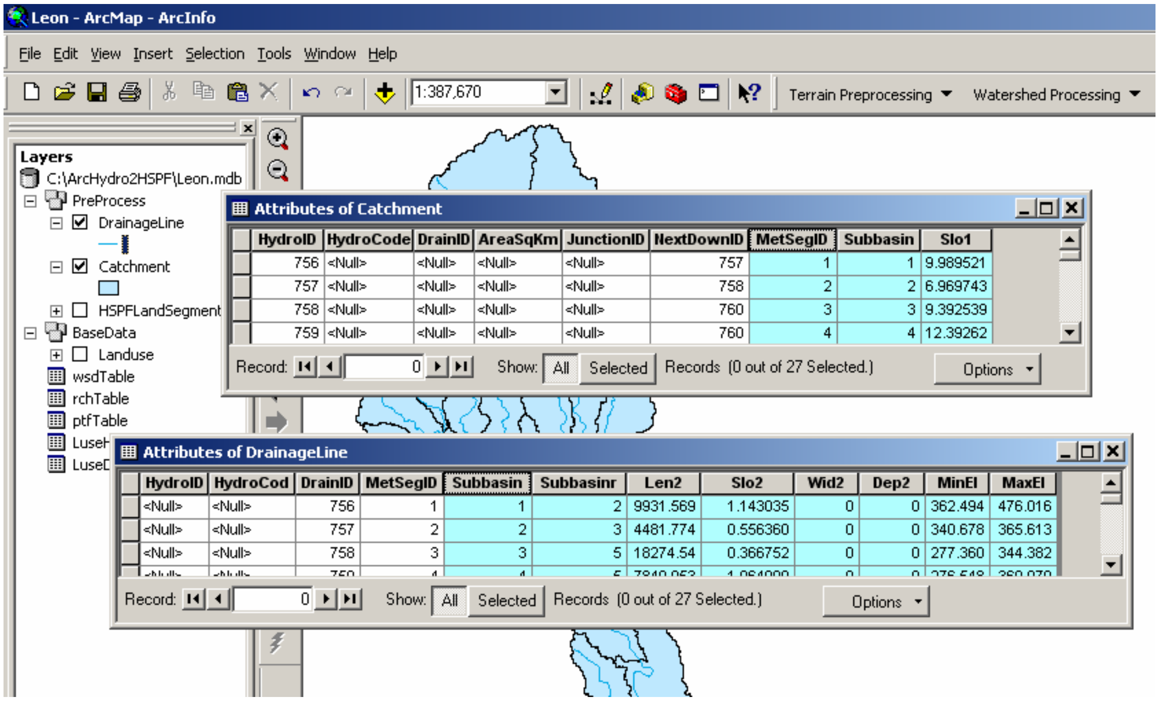


Figure 6.4 Feature classes after calculating attributes for HSPF model development.

Polygon landuse data from the USGS's GIRAS dataset was used for three of the projects, and the remaining two used raster data from the National Land Cover Dataset (NLCD). Figure 6.5 shows the structure of the HSPF Preprocessing geodatabase, Arc Hydro Catchment and DrainageLine data, and Landuse data for Leon Creek.

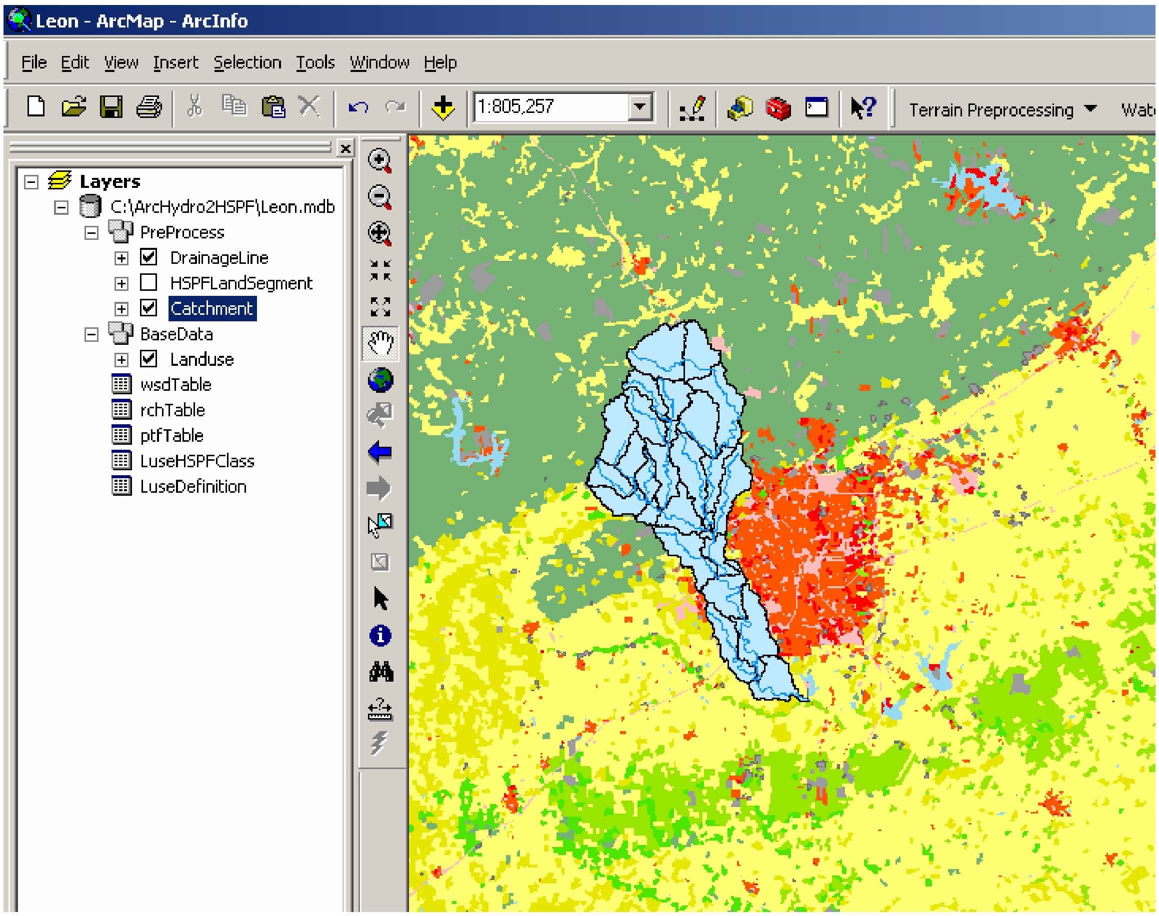


Figure 6.5 Arc Hydro and landuse data in ArcGIS HSPF Preprocessing geodatabase

Catchment and Landuse data are combined using the ArcGIS HSPF Preprocessing methodology to define land areas to be simulated with HSPF. The result is a feature class with a structure very similar to that of the .wsd file used to transfer data to WinHSPF. Figure 6.6 shows this feature class in which each feature has attributes defining what type of landuse it represents, what subbasin it belongs to, and the average slope of the subbasin. Three of these features are highlighted in the table shown in Figure 6.6 and their geospatial representations are also highlighted in the background map.

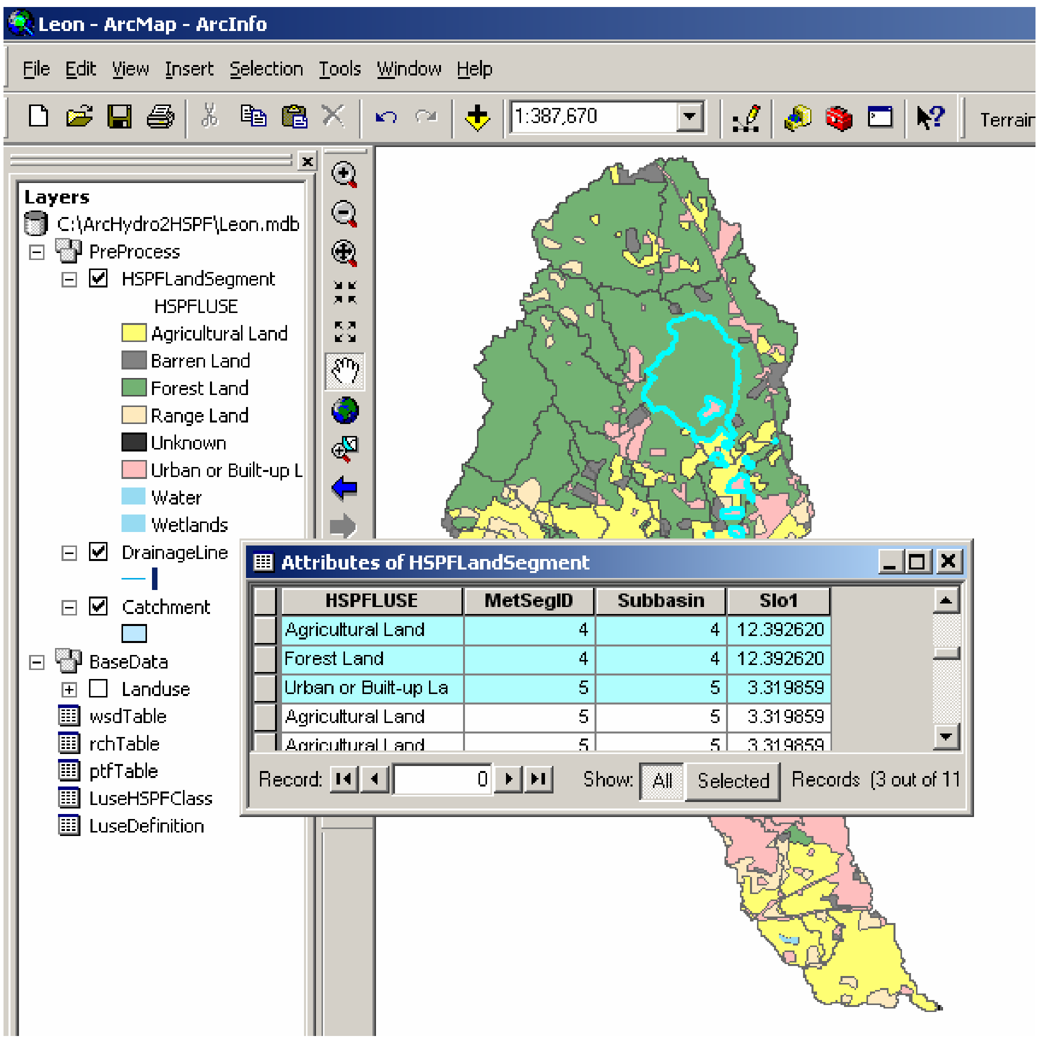


Figure 6.6 HSPFLandSegment feature class. Used to write .wsd file defining land segments to be simulated by HSPF.

Information from this Land Segment feature class and the DrainageLine feature class are written to intermediate text files using custom ArcGIS tools. Figure 6.7 shows the four intermediate text files, .wsd, .rch, .ptf, and .psr which are required to build a new .uci file with the WinHSPF system.

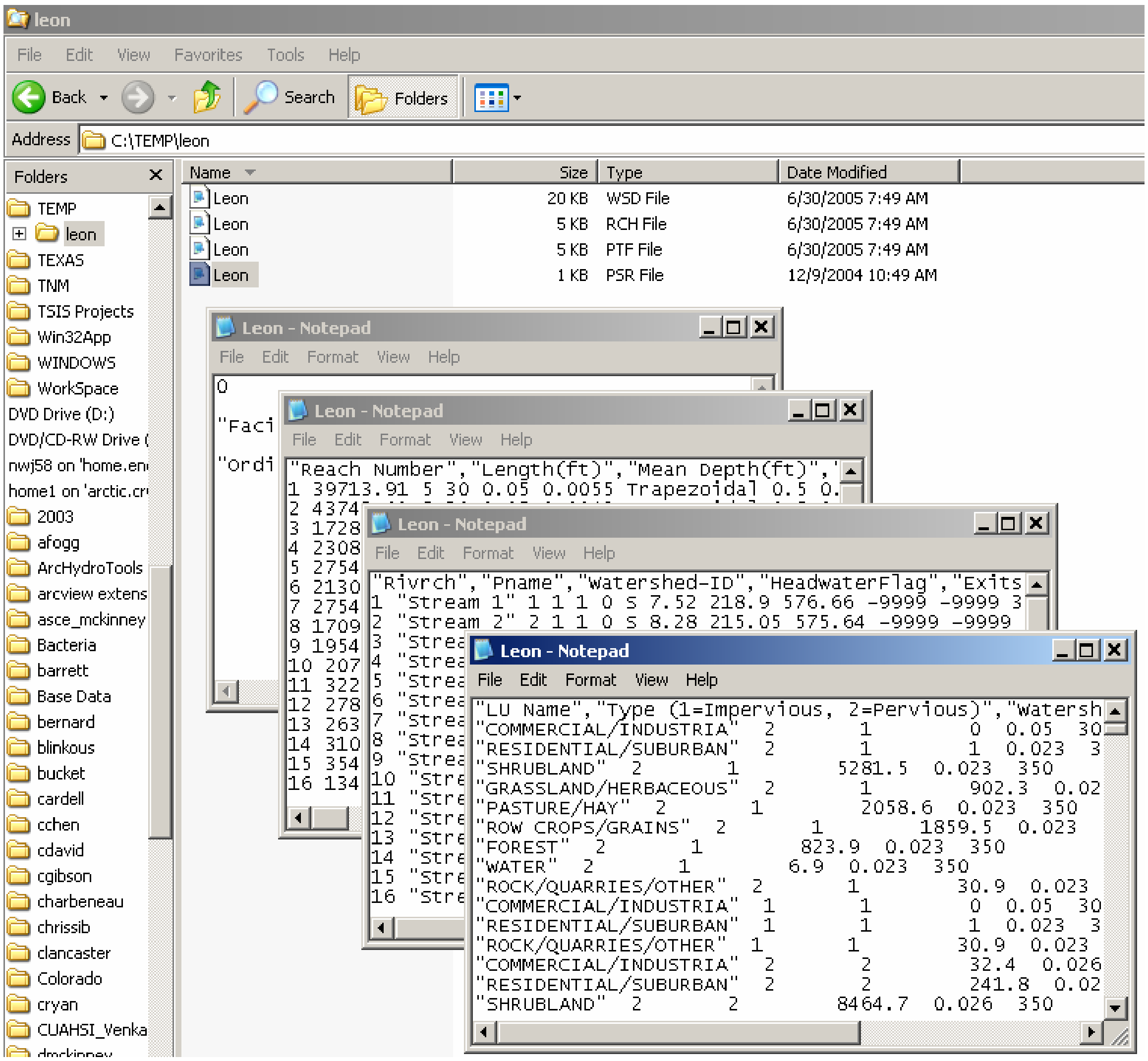


Figure 6.7 Intermediate text files.

The intermediate text files shown in Figure 6.7 are used to build a new .uci file using the “Create Project” tool of the WinHSPF program. The .uci file resulting from this tool is shown in Figure 6.8.

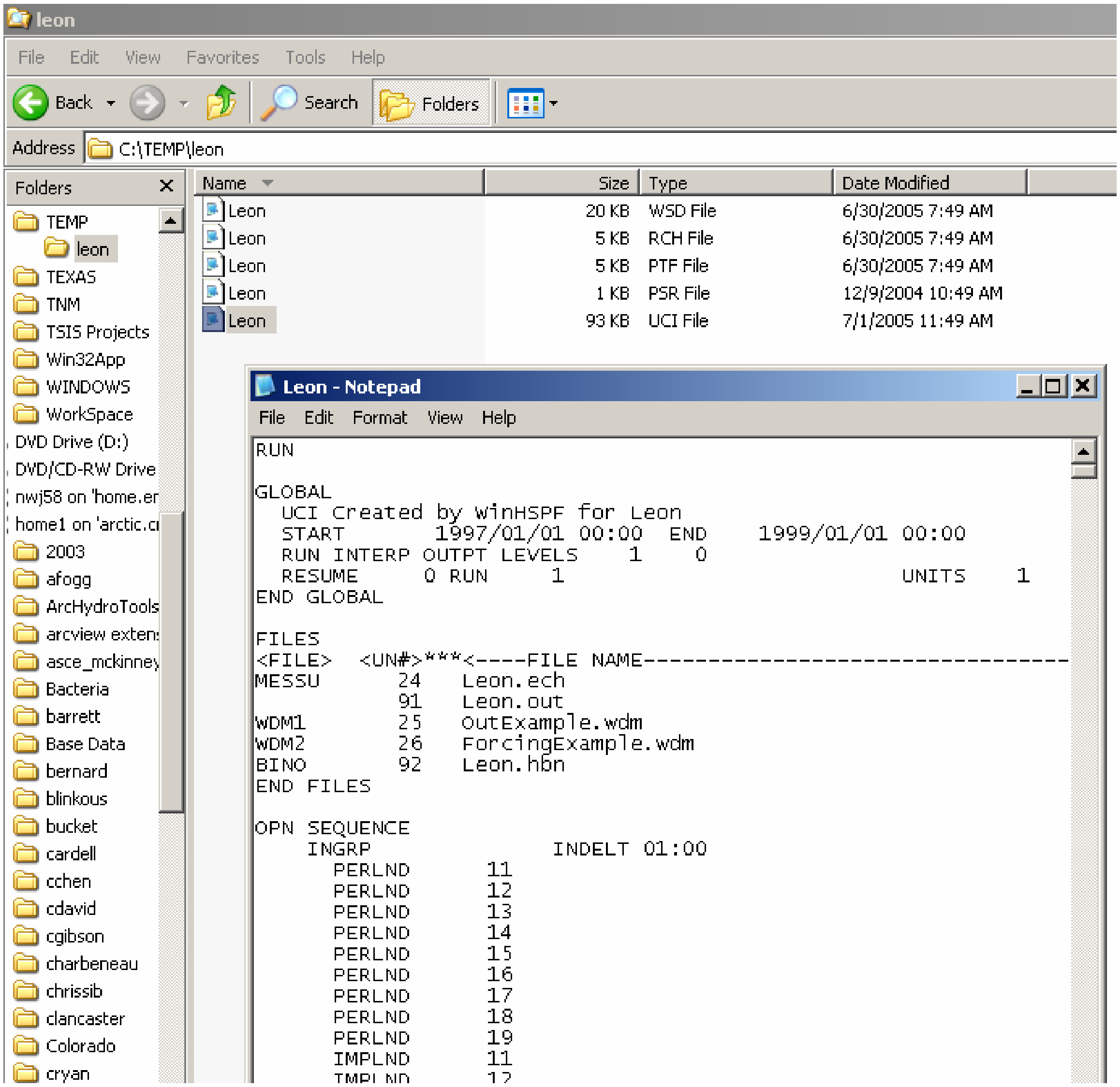


Figure 6.8 New .uci file created with ArcGIS HSPF Preprocessing methodology.

The steps outlined above were used to develop HSPF model files for 5 study areas in the ArcGIS environment. These models could have been developed using the BASINS ArcView 3.x system, however, at the Center for Research and Water Resources, HSPF users are more familiar with ArcGIS and Arc Hydro than with the ArcView 3.x BASINS system. Arc Hydro GIS data provides a good starting point for HSPF model development for those familiar with ArcGIS and also facilitates the use of ArcGIS tools and data structures after initial model development.

The models could also have been developed by manually building a new .uci file using data developed in ArcGIS. This would be tedious and error prone and the ArcGIS HSPF Preprocessing methodology streamlines the task considerably.

6.2 Results of ArcGIS Timeseries Preprocessing Applications

For three of the projects presented above, ArcGIS tools and data structures have been utilized after initial model development to prepare input timeseries for HSPF modeling. The ArcGIS Timeseries Preprocessing methodology assumes a standard HSPF model configuration so that GIS data can be used to transfer information to model files after initial model development. Input time series data, developed using NEXRAD estimates for rainfall and stored in the Arc Hydro time series format, is used to write the required timeseries files and to update model files to read from the appropriate timeseries dataset.

Estimates for basin-average-precipitation were developed with NEXRAD data using GIS tools and data structures. NEXRAD precipitation data is typically stored in a format that is not compatible with the .wdm file format required for HSPF modeling, however, the Arc Hydro time series format provides a data structure which can reconcile the differences. The details of the process used to develop basin-average-precipitation are outlined in Appendix B. Figure 6.9 illustrates the NEXRAD cells and the GIS data used to develop basin-average precipitation for the Leon Creek watershed.

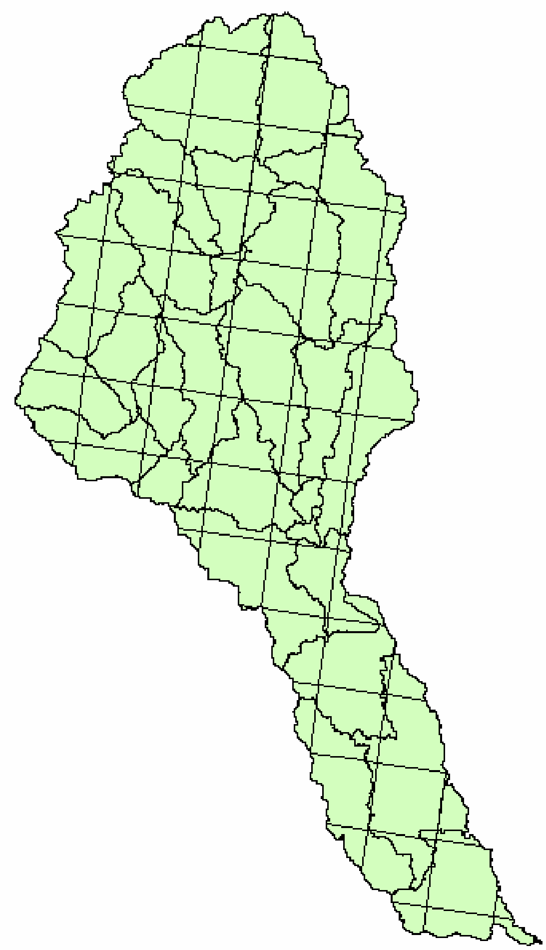


Figure 6.9 NEXRAD data overlain on GIS MetSegments.

The central data structure to the ArcGIS Timeseries Preprocessing methodology is the GIS MetSegment feature class. Each GIS MetSegment feature is attached to Arc Hydro time series data using the standard Arc Hydro structure. Figure 6.10 shows the attributes used to link GIS MetSegments to Timeseries values.

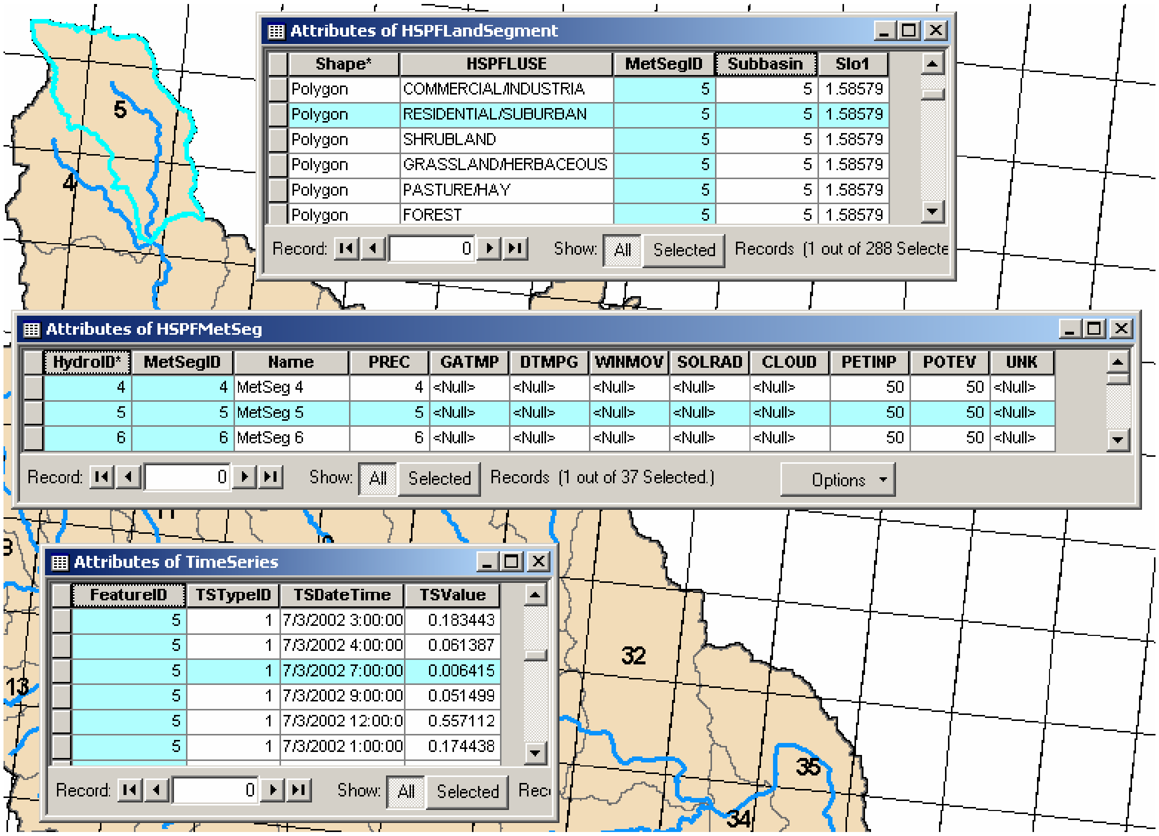


Figure 6.10 GIS MetSegment feature class: Example of Application

ArcGIS Timeseries Preprocessing tools were first used to write time series data from the Arc Hydro time series structure to .wdm files for input to the HSPF model. The .wdm dataset locations were specified in attributes of the HSPFMetSeg feature class.

In the ArcGIS Timeseries Preprocessing application used to develop input data for the Attascosa HSPF model, GIS MetSegment boundaries followed the drainage area boundaries of each River Segment. In this way, basin-average-precipitation estimates developed from NEXRAD data were applied to the HSPF model.

Figure 6.11 shows the .wdm file resulting from the application of a custom ArcGIS Geoprocessing tool to write precipitation data from the Arc Hydro geodatabase to a .wdm file. The file is shown in the WDMUtil program and represents the input time series for the Atascosa HSPF model.

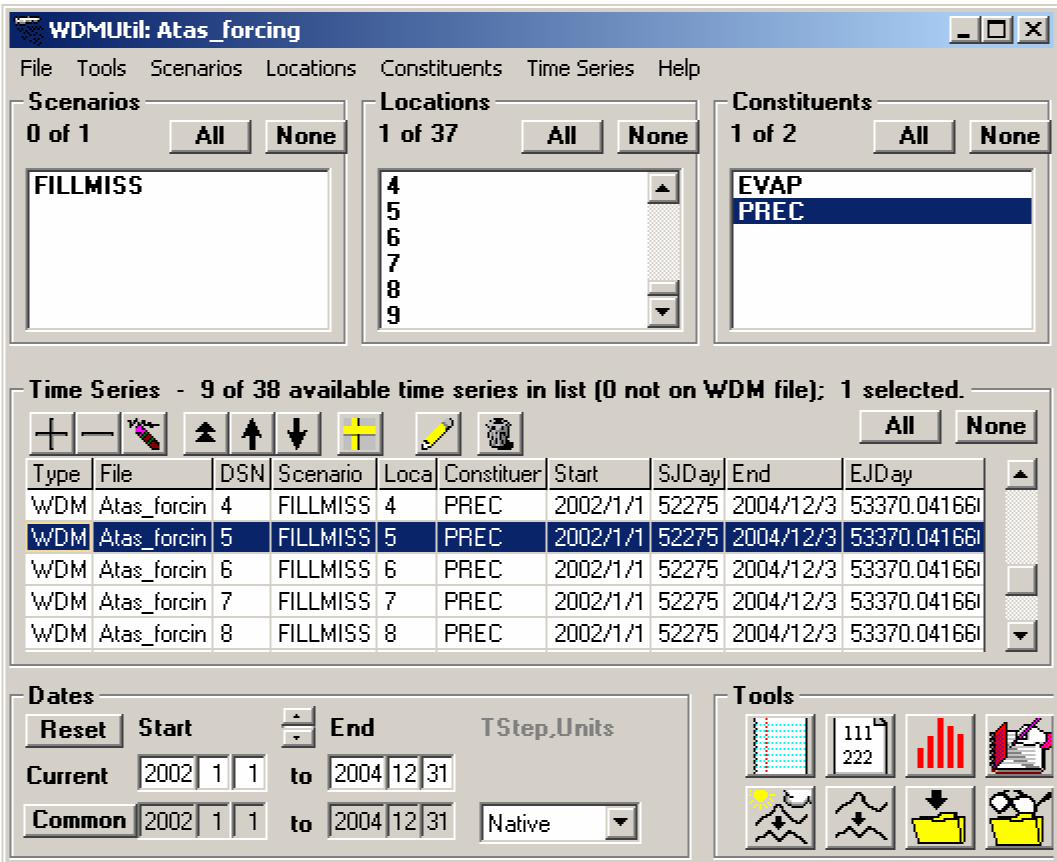


Figure 6.11 Results of ArcGIS Timeseries Preprocessing system: .wdm file.

After the input time series was written to the appropriate .wdm datasets, the HSPF .uci file was updated using the GIS representation of the Model Elements. Appropriate lines were added to the .uci file specifying the input precipitation time series for each Model Element. The updated External Sources block of the .uci file, HSPFLandSegment feature class and MetSegment feature class are shown in Figure 6.12.

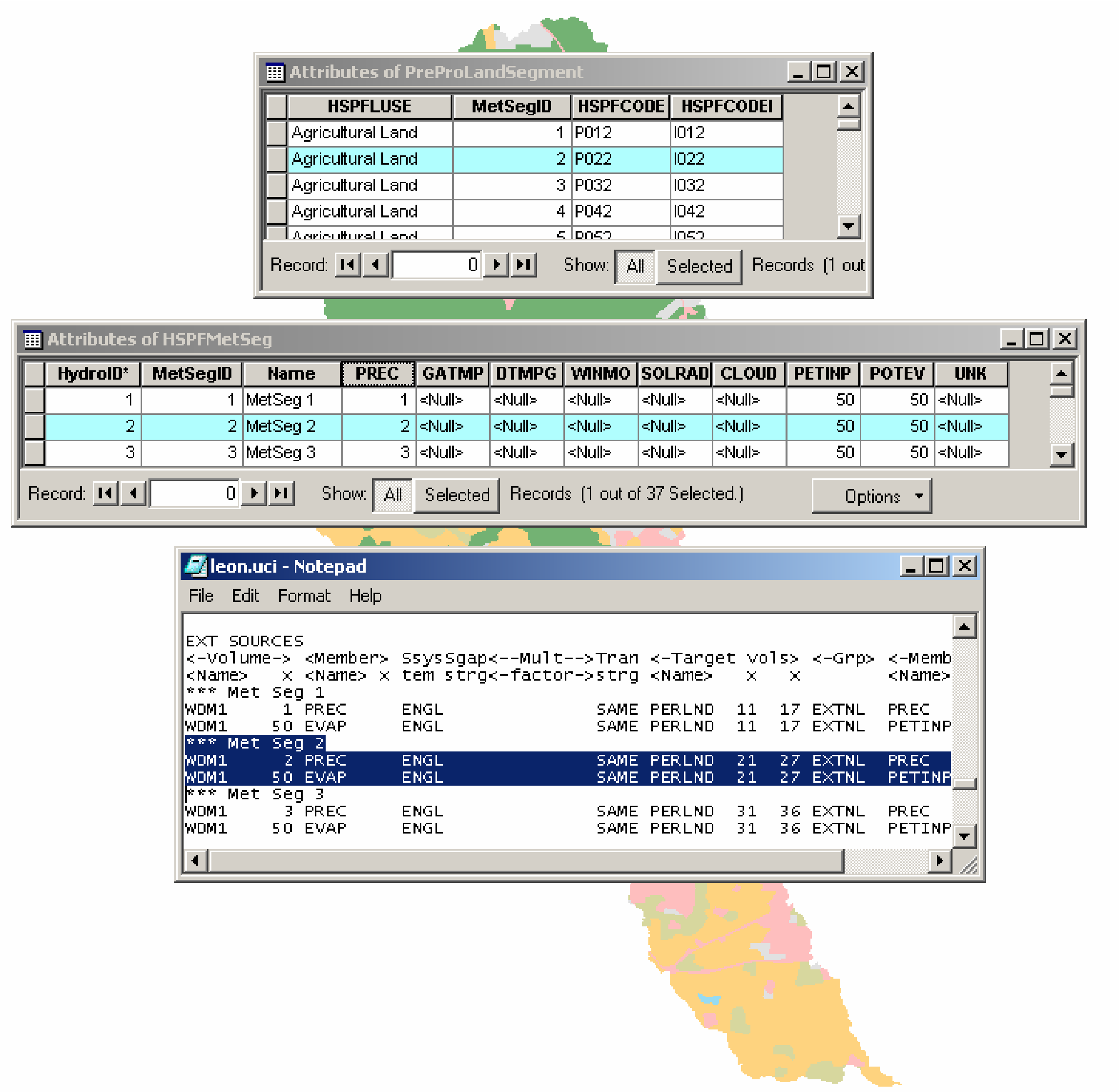


Figure 6.12 Results of ArcGIS Timeseries Preprocessing system: Updated EXTSOURCES block.

The transfer of information from GIS data (HSPFMetSegment feature class and Arc Hydro time series) and the HSPF Model Elements (.uci file) was facilitated by the presence of a unique identifier in both the geodatabase and .uci file. The HSPFCode is a string of characters that is stored both on the HSPF Model Element in the .uci file and on the feature in the geodatabase that represents its location in the real world. Tools from the ArcGIS HSPF Preprocessing system were used to assign the HSPFCode to the HSPF .uci file and the GIS data used in its development. Figure 6.13 shows the HSPFCode in both the .uci file and the GIS LandSegment data.

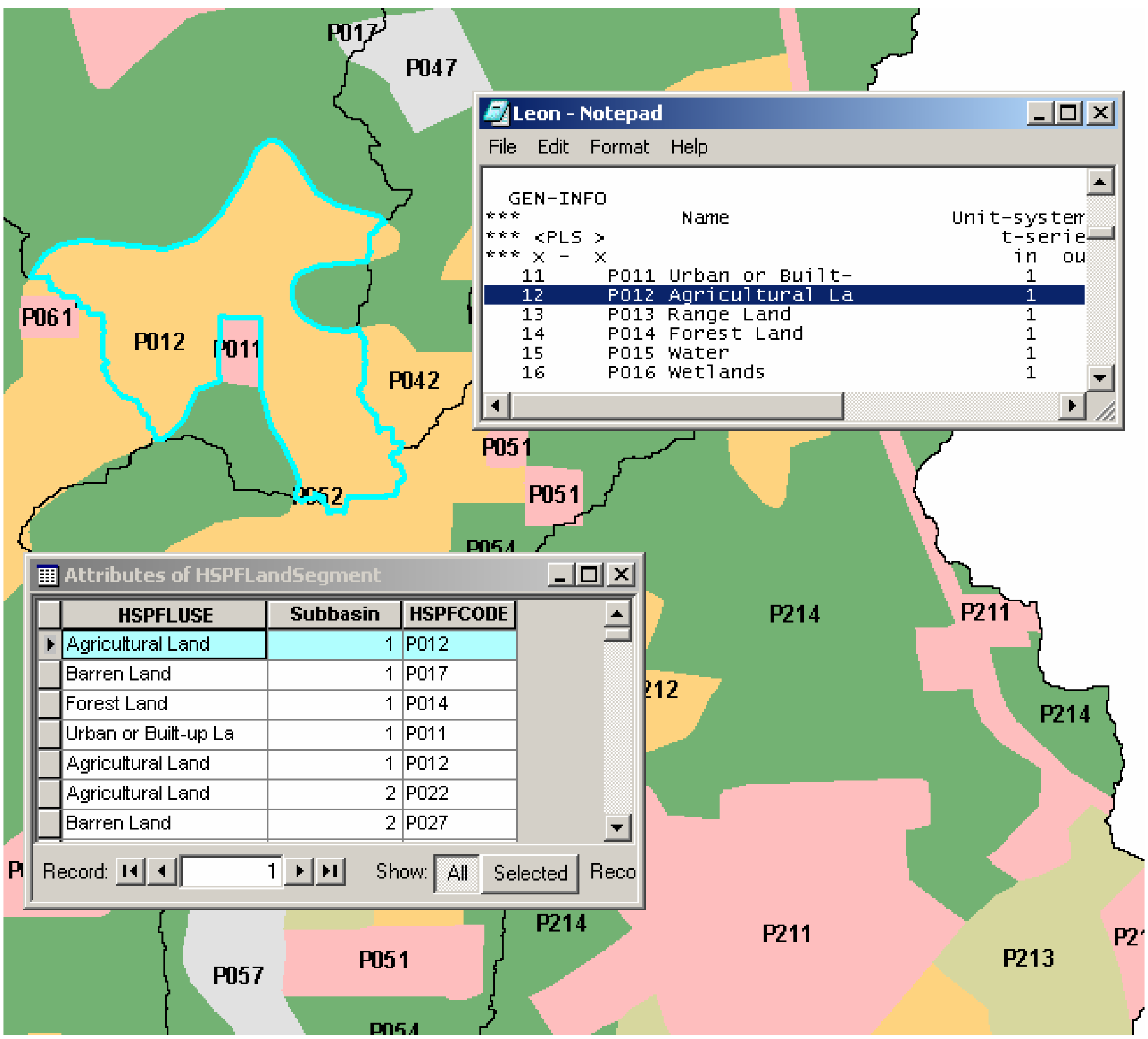


Figure 6.13 HSPFCode used to implement ArcGIS Timeseries Preprocessing system.

The steps outlined above were used to develop input time series for three separate HSPF models. In each case, NEXRAD time series data were developed using GIS and applied to HSPF model files using the GIS representation of the model elements. The ability transfer of information between GIS data and HSPF model files is a result of an attempt to maintain a geospatial representation of Model Elements after initial model development. This would not be possible with the BASINS HSPF Preprocessing system because all the GIS functionality is hidden behind a custom tool which does not retain any GIS data during model development.

6.3 LIMITATIONS

A methodology for developing HSPF models in the ArcGIS environment has been implemented with very little custom tool development; however, its use of existing tools comes at a cost. The methodology is not as clean as the one-button type implementation of the BASINS system. The most significant difference between the BASINS HSPF Preprocessing system and the ArcGIS HSPF Preprocessing system is the exposure of many of the functions that were hidden behind the user interface in the ArcView 3.x system. In the BASINS ArcView 3.x system, once physical parameters were defined in GIS data, a single tool was used to define land segments, write intermediate text files, and build a new .uci file using WinHSPF.

The tools of the ArcGIS HSPF Preprocessing system can be linked together to provide a one-button-type application using ESRI's ModelBuilder, however, the system is not as robust as the ArcVeiw 3.x implementation in the BASINS system. Beginning ArcGIS users will likely find the methodology complicated compared to the ArcView 3.x BASINS system; however, advanced users may find some parts of the system more flexible and the data management framework more robust.

While the tools and data structures used in the ArcGIS HSPF Preprocessing system do provide a functional way of developing HSPF models in the ArcGIS environment, they are not a “Production-Level” toolset and many improvements could be made. One flaw in the design is the necessity of first extracting information from GIS data to tables in the database and subsequently writing this data to text files. A much more efficient method would write the data directly from GIS data to text files.

An additional complication involves the storage of network information in both the Arc Hydro attributes and those used by the BASINS 3.x system. While this redundancy is inefficient and unnecessary, an attempt was made to keep the ArcGIS HSPF Preprocessing geodatabase structure as consistent with the previous ArcView 3.x file structure as possible. As such, attributes of feature classes in the ArcGIS system have the same name as their ArcView 3.x shapefile counterparts. In this way, shapefile data developed using the BASINS ArcView 3.x system can be directly loaded into the ArcGIS database without having to navigate through differences in naming conventions.

The ArcGIS Timeseries Preprocessing methodology is intended to be a demonstration of how ArcGIS can be used even after initial HSPF model development. The system was designed primarily to use NEXRAD precipitation data to drive HSPF models. The configuration convention required to implement the process is convenient for applying different precipitation timeseries to different parts of an HSPF model but creates very large model files. With a unique set of land segments simulating the drainage area for each individual stream segment, there could be as many as 10-15 land segments for each river segment in the model. If more than 30 stream segments are to be simulated, the number of HSPF Model Operations could easily exceed the limit of 500 Operations in a single HSPF run. Though some parts of the system may be helpful for specific applications, the ArcGIS Timeseries Preprocessing system is not necessary for many HSPF modeling objectives.

The HSPFCode concept is a general, robust way of transferring information between GIS data and the .uci file, however, the tools used to assign the HSPFCode are not well developed. The convention used in this research for the HSPFCode is based on the Operation Numbering convention adopted by WinHSPF when building a new .uci file, and only works well for models with ten or more River Segments.

To some, the most significant limitation of the ArcGIS HSPF Preprocessing methodology may be the continued use of proprietary GIS software. Though its analytical and data management capabilities are unmatched by other GIS, ESRI software and licenses are extremely expensive and prohibitive for many HSPF modeling applications. The core component of the newest release of the BASINS system, v.4.0, will not be tied to any proprietary software. Certain BASINS features will still be dependent upon proprietary GIS packages, but with BASINS 4.0 a user will be able to download data, delineate watersheds, and set up HSPF without purchasing GIS software. (Duda, unpublished). The continual updating of the BASINS system to be compatible with new releases of ESRI products is inefficient for EPA funding and requires the users of BASINS to continue to purchase proprietary software products.

The developers of BASINS are able to break the ties to proprietary GIS packages because of the current availability of open source GIS packages. Though not as well developed as ESRI products, some open source GIS libraries can read GIS data from several different formats and provide many of the spatial analysis capabilities necessary for HSPF model preprocessing. Though far from providing the complex spatial analysis and data management framework afforded by ESRI products, the development of open source GIS may someday provide an alternative for the spatial analysis required for hydrologic model development.

6.4 BROADER CONTEXT: GIS AND HYDROLOGIC MODELING

Though GIS tools from the BASINS system provide a well-structured, convenient way of using GIS data to develop HSPF models, many GIS users are moving away from the ArcView 3.x software in favor of the newer ArcGIS products. The use of ArcGIS in the research community is growing rapidly because of the ease with wich custom applications can be developed. ESRI no longer supports the Avenue and ArcInfo libraries used to develop custom tools in ArcView, and the COM-Compliant ArcObjects used by ArcGIS provide a welcome alternative for many developers.

The ArcGIS HSPF Preprocessing methodology essentially mirrors the procedure used by the BASINS ArcView 3.x HSPF Preprocessing system, but implements it in the ArcGIS environment. The tools of the ArcGIS HSPF Preprocessing methodology provide users of ArcGIS with the tools necessary to efficiently produce HSPF models in a similar manner as that provided by the BASINS ArcView 3.x system. It is impractical for GIS users familiar with ArcGIS products to continue to continue using ArcView software only for HSPF preprocessing when ArcGIS has all of the required spatial analysis capabilities and a much more robust data management framework. HSPF models could be developed on a model-by-model basis in the ArcGIS environment; however, the methodology developed here provides a well-structured, reproducible way of using ArcGIS to develop HSPF models.

One of the reasons why NEXRAD data is not widely used for hydrologic modeling is questions about its accuracy. No research to date has shown that significantly better results are obtained from hydrologic models using NEXRAD estimates instead of conventional gauge estimates (Nearey, et al. 2004). The ArcGIS Timeseries Preprocessing methodology developed in this research provides the tools necessary to conduct studies comparing the two data sources using the HSPF model.

Though accuracy questions are a major obstacle to the use of NEXRAD in hydrologic modeling, another difficulty is the vast differences in file format. When building a temporally continuous timeseries (.wdm file for HSPF modeling) from a collection of spatially continuous data sets (what NEXRAD data is stored in) it may be necessary to open thousands of gridded files and pick out only a few values from each. If the data is to be stored in .wdm files for HSPF modeling, each value retrieved from a gridded data set must be placed in its own timeseries data set corresponding to its spatial location.

The ability of the Arc Hydro timeseries structure to store both types of timeseries and efficiently prepare the data for HSPF modeling demonstrates the power of ArcGIS data management capabilities. The ArcGIS Timeseries Preprocessing methodology provides a solution for reconciling the differences between typical NEXRAD and .wdm file formats and provides a way to apply the data directly to HSPF models. Though the question of NEXRAD accuracy remains to be addressed before radar data will be widely used in hydrologic modeling, ArcGIS data structures and tools are capable of providing a method for transferring data between the two very different file formats.

Chapter 7 Conclusions

The overall objective of this research is to bridge the gap that exists between ArcGIS and the HSPF model, thereby placing the HSPF model in the broader context of a data management framework and supporting set of tools that is widely used in the hydrologic modeling community. This is accomplished through the development of a preprocessing methodology that makes use of the capabilities of existing GIS and non-GIS tools. The ArcGIS HSPF Preprocessing methodology implements almost exactly the same algorithms as the HSPF-related BASINS tools but implements the process in the ArcGIS environment using a more robust data management framework.

In order to demonstrate the ability of ArcGIS data structures and tools to aid HSPF modeling after initial development, a methodology is also presented to organize and prepare input time series data using the Arc Hydro time series structure. The ArcGIS HSPF Timeseries Preprocessing methodology uses rainfall data stored in a well-structured relational database format to prepare .wdm files for HSPF modeling using custom ArcGIS tools. In addition to writing the necessary time series files, custom ArcGIS tools automatically update HSPF model input files to receive precipitation data from appropriate datasets based on information stored in the database. The ArcGIS Timeseries Preprocessing methodology provides one example of how ArcGIS data structures and tools can be used to aid in HSPF modeling after initial model development.

7.1 ArcGIS and Initial HSPF Model Development

The main contribution of this research is the development of a standard methodology for preparing HSPF models in the ArcGIS environment. For HSPF modelers who wish to use ArcGIS in lieu of ArcView 3.x, the ArcGIS HSPF Preprocessing methodology provides a practical way of developing new HSPF models from GIS data. By linking existing tools from the Arc Hydro and BASINS systems with a few custom ArcGIS tools, the ArcGIS HSPF Preprocessing methodology avoids much duplication of effort and provides a new method for developing HSPF models in the ArcGIS environment.

The Arc Hydro data model and terrain processing tools are widely used in the hydrologic modeling community for defining and organizing basic data in the ArcGIS environment. The ArcGIS HSPF Preprocessing methodology makes use of this well-developed data structure and avoids the development of an additional ArcGIS terrain processing toolset by starting with data from the Arc Hydro data model. Though the ArcGIS HSPF Preprocessing methodology is designed to begin with data from the Arc Hydro data model, the tools are flexible enough and the methodology general enough to be applied to data from other sources as well.

In addition to making use of the existing Arc Hydro tools and data model, the ArcGIS HSPF Preprocessing methodology also makes use of WinHSPF, a non-GIS component of the BASINS system. WinHSPF is most often used in initial HSPF model development following spatial analysis with the ArcView 3.x components of the BASINS system. However, because its algorithms require no complex spatial analysis, WinHSPF can be used to build a new .uci file independently of the GIS components of the BASINS system if the required data is available from another source. The ArcGIS HSPF Preprocessing methodology is designed to make use of the capabilities of WinHSPF (and the substantial effort that went into its development) by preparing ArcGIS data to provide the information required to build a new HSPF model. WinHSPF is an extremely important tool in the HSPF modeling community and its inclusion in the ArcGIS HSPF Preprocessing methodology provides a well-established component that will be familiar to many users.

The ArcGIS HSPF Preprocessing Methodology does require the use of proprietary ArcGIS software and data structures. This requirement may be a limitation for some users because of the cost of ArcGIS software and licenses. For those already familiar with ArcGIS, the methodology developed here provides an accessible, efficient, reproducible way to use ArcGIS for HSPF model development and provides a starting point for using ArcGIS after initial model development.

7.2 Use of ArcGIS in HSPF Time Series Development

The aforementioned methodology for using ArcGIS data and tools to develop initial HSPF models is only the beginning of the possibilities for utilizing ArcGIS in HSPF modeling. The ArcGIS Timeseries Preprocessing methodology provides one example of how using ArcGIS data for HSPF modeling places the HSPF model in the powerful spatial analysis and data management framework provided by ArcGIS.

NEXRAD data is commonly stored in a file format that is entirely incompatible with the .wdm file format used for HSPF modeling. However, with NEXRAD data stored in the relational database structure used by the Arc Hydro timeseries format, it can easily be extracted and written to the files necessary for HSPF modeling. In addition to providing access to NEXRAD data, the ArcGIS Timeseries Preprocessing methodology demonstrates how HSPF model files can be linked to GIS data. With a geospatial representation of the Model Elements simulated by HSPF, information stored in GIS data can be used to automatically update model files.

ArcGIS and Arc Hydro are powerful tools for applications to hydrologic and water quality modeling because they provide general way of representing the the environment through which water flows that is independent of any specific application. The fact that NEXRAD data, which is stored in an entirely different format than .wdm time series required for HSPF modeling, can be used to develop input time series for HSPF modeling using GIS data and tools demonstrates this power. Many hydrologic data and model files are far removed from each other because of incompatibilities in file format or data structure. All hydrologic data and models, however, describe the movement of water through the same physical environment, and GIS can be used to facilitate the transfer of information between otherwise incompatible models and data formats by providing an accurate, consistent, widely-used abstraction of that environment.

7.3 Future Work: Broader Implications for the Use of ArcGIS in HSPF Modeling

In addition to the development of input precipitation data, ArcGIS tools provide a host of other capabilities that could be useful for HSPF modeling. ArcGIS is widely used for visualization of data and many basic tools are included in standard ESRI software packages. The ArcGIS Timeseries Preprocessing methodology only involves the transfer of information from GIS data to HSPF model files, but the reverse process could allow the results of HSPF simulations to be visualized using the powerful tools and data management framework provided by ArcGIS. This process has already been implemented for the HEC-RAS hydraulic models to visualize flood inundation polygons using ESRI's ArcMap software (Roboyo 2004), and the results of HSPF simulations could be similarly viewed using ESRI's tracking analyst. ESRI is working to include more time series capabilities in their newest release (ArcGIS 9.2), and these additions will provide more possibilities for storing, managing, and visualizing HSPF time series data in ArcGIS.

Neither the ArcGIS HSPF Preprocessing methodology nor the BASINS system contain extensive tools to deal with river channel geometry, however, tools for developing detailed descriptions of river channel geometry using Arc Hydro and ArcGIS tools have been developed at CRWR (Mirwade 2004). The River Channel Morphology Model (RCMM) uses average channel width and depths at known locations and a GIS description of the sinuosity of the river channel to develop a mathematical description of river channel geometry. RCMM uses ArcGIS data to develop a digital description of river channel geometry at locations where detailed cross section information is unavailable. With an ArcGIS framework for developing HSPF models, this work could be incorporated into model development to create more accurate descriptions of river channel geometry for HSPF modeling.

In addition to the analytical capabilities of ArcGIS, its data management framework provides a robust structure for storing geospatial and temporal data. The Bexar Regional Watershed Management Coalition's (BRWMC) Regional Watershed Management plan is built around a digital description of the hydrologic environment using the Arc Hydro framework. The HEC-HMS and HEC-RAS models have already been linked through ArcGIS and Arc Hydro data (Roboyo 2004), and this research is part of a larger effort to incorporate HSPF into the system. Hydrology (HEC-HMS) and water quality (HSPF) models will be built using the same drainage areas defined in GIS data and could potentially share the same input timeseries data. Eventually, a database containing NEXRAD rainfall data could be used to develop input time series for both hydrologic and water quality models using the spatial analysis and data management capabilities of ArcGIS.

Because both hydraulic and water quality models are built upon the same GIS data, detailed descriptions of channel geometry used for hydraulic modeling will be also be available for use in HSPF model development. The foundation of the Watershed Management Plan is GIS data containing the most accurate, up-to-date, digital description of the environment. Models for hydrologic, water quality, and other processes will all be built upon this GIS description and therefore reflect the best information available about the environment.

Appendix A: Using Raster Landuse Data

Contents:

This appendix contains a description of a method for using raster landuse data in the ArcGIS HSPF Preprocessing methodology.

Objective:

The objective of the following tasks is to create a feature class in which each feature represents the area of a single type of landuse that contributes to a single River Segment. This data will play the role of the HSPFLandSegment feature class described in Chapter 5.

Tasks:

Four major tasks will be accomplished to use raster landuse data with the ArcGIS HSPF preprocessing methodology.

(1) Prepare landuse data
(2) Calculate contributing area for each River Segment (using Catchment feature class)
(3) Copy tabulated areas to Catchment feature class
(4) Load data into HSPFLandSegment feature class

Specific Steps:

After the slope attributes are calculated for the DrainageLine and Catchment feature classes:

<u>*prepare landuse data*</u>

(1) populate LuseHSPFClass table to define the 9 or fewer types of land to be simulated by HSPF and the impervious area associated with each type
(2) resample raster data to 9 or fewer types defined in LuseHSPFClass table using the “Reclass” (manual) or “Reclass by Table” (automatic) in the “3D Analyst – Raster Reclass” toolbox.

<u>*calculate contributing areas*</u>

(3) use ArcToolbox "Tabulate Area" function (spatial analyist extension required) in the "Spatial Analyst Tools – Zonal" toolbox to tabulate the areas of each type of landuse in the Catchments. Tabulate area based on a unique number in the Catchment feature class (Subbasin or HydroID) and save the table to a convenient location

copy tabulated areas to Catchment

(4) Open the attributes of the Catchment feature class and add two attributes for each type of landuse found in the 'Tabulate Area' function. One attribute is a text attribute and will store the string (HSPFLUSE) from the HSPFLuseClass table. The other is of type 'double' and will store the area associated with this landuse type. Give them meaningful names like "lu1", "a1," "lu2," "a2," etc…

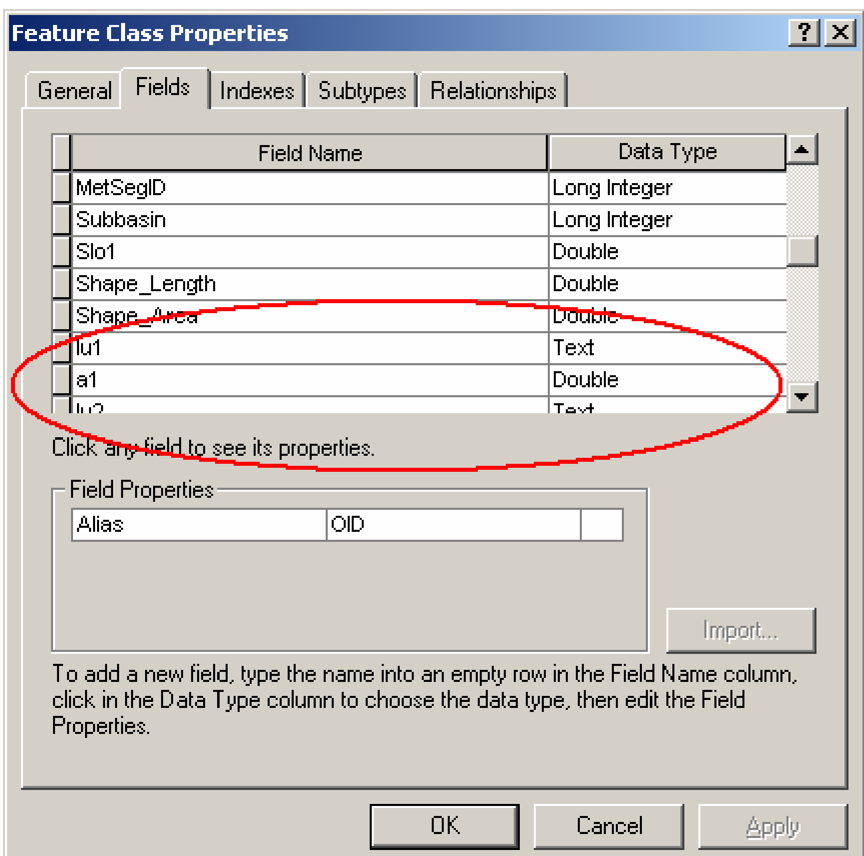


Figure A.1 Add attributes to store landuse types and areas

(5) Open the new table and the Catchment feature class in ArcMap and join them based on the unique ID used to "Tabulate Area".

(6) Copy the areas from the table to the appropriate (new) columns in the Catchment feature class. *Note: 5 and 6 could equivalently be accomplished using the "Copy Field: Table to Feature Class" tool repeatedly either through the GUI or from the command line.*

(7) Populate the attributes of the Catchment feature class defining the types of landuse to be simulated with the text descriptions from the LuseHSPFClass table. This can be done by right clicking on the field and choosing 'Calculate Field.' It is recommended to use the copy-paste function to ensure the exact same string is present in the attributes of Catchment.

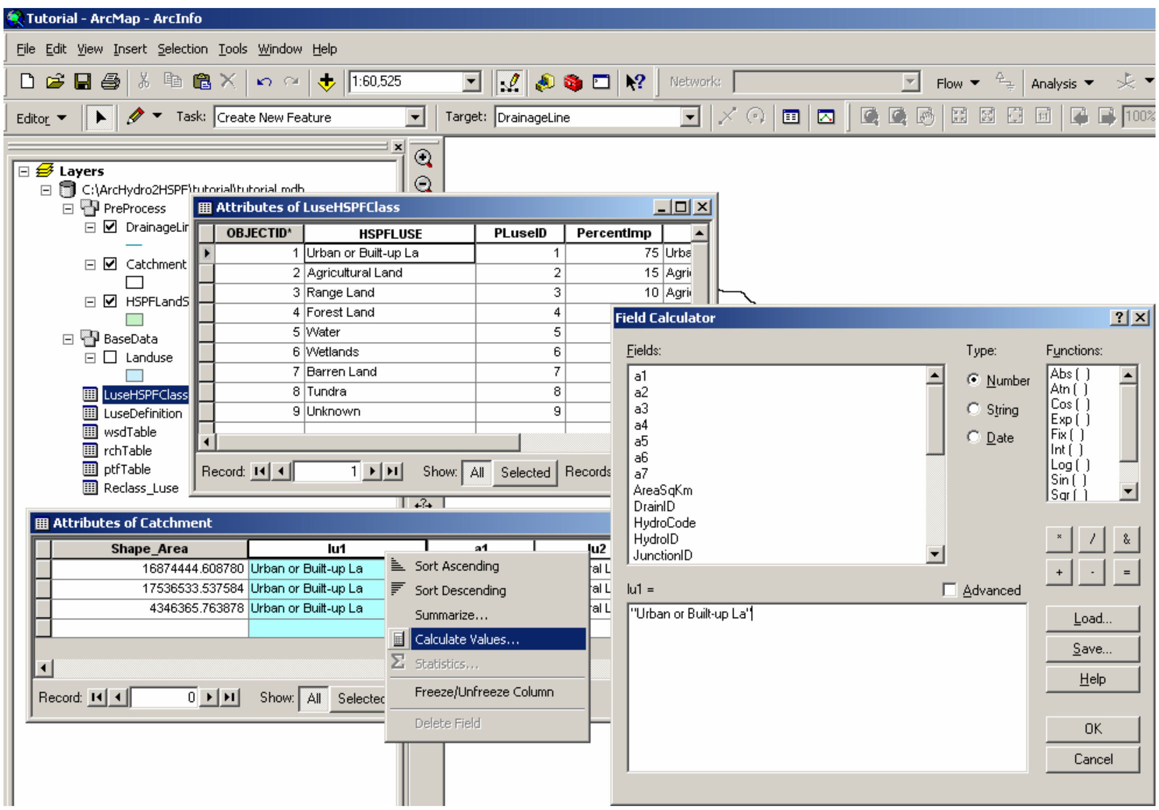


Figure A.2 Calculate attributes decribing HSPF landuse types in ArcMap

<u>*load data into HSPFLandSegment feature class*</u>

(8) The data is now ready to be loaded into the HSPFLandSegment feature class. In ArcCatalog, right click on the HSPFLandSegment feature class and choose 'Load Data'

(9) Browse to the ‘Catchment’ Feature Class, add it to the list and click next twice to get to the field mapping screen

(10) Attributes for HydroID, MetSegID, Subbasin, and Slo1 should be automatically selected. For the ‘HSPFLUSE’ attribute, choose the text field for the first type of landuse to be simulated by HSPF. For the ‘EffArea’ field choose the area corresponding to the first type of landuse to be simulated by HSPF. Click Next to finish loading the data.

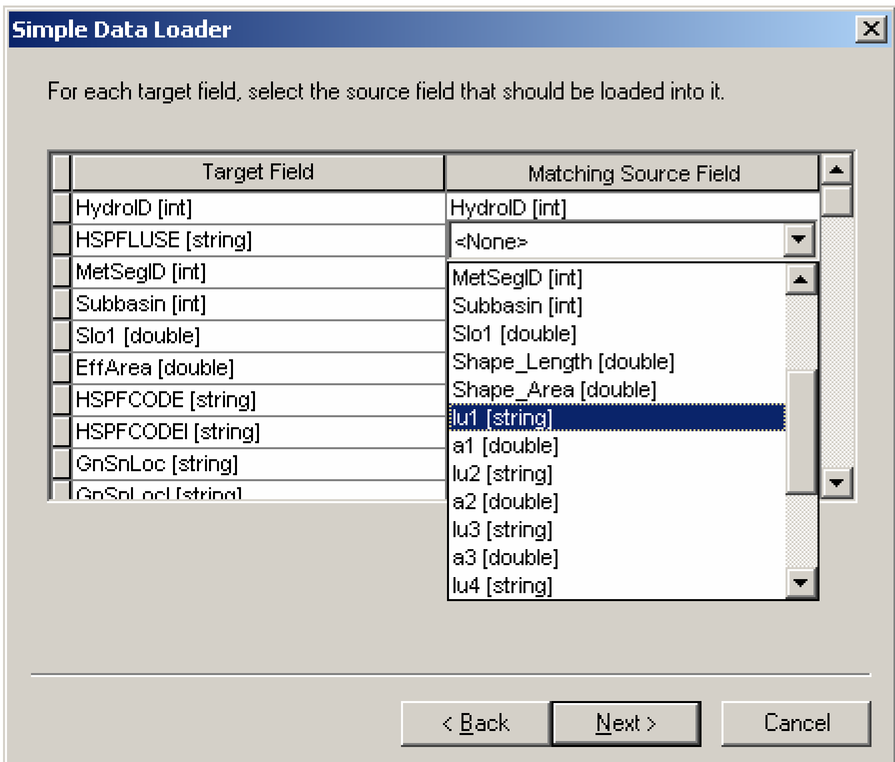


Figure A.3 Load data into the HSPFLandSegment feature class: Landuse Types

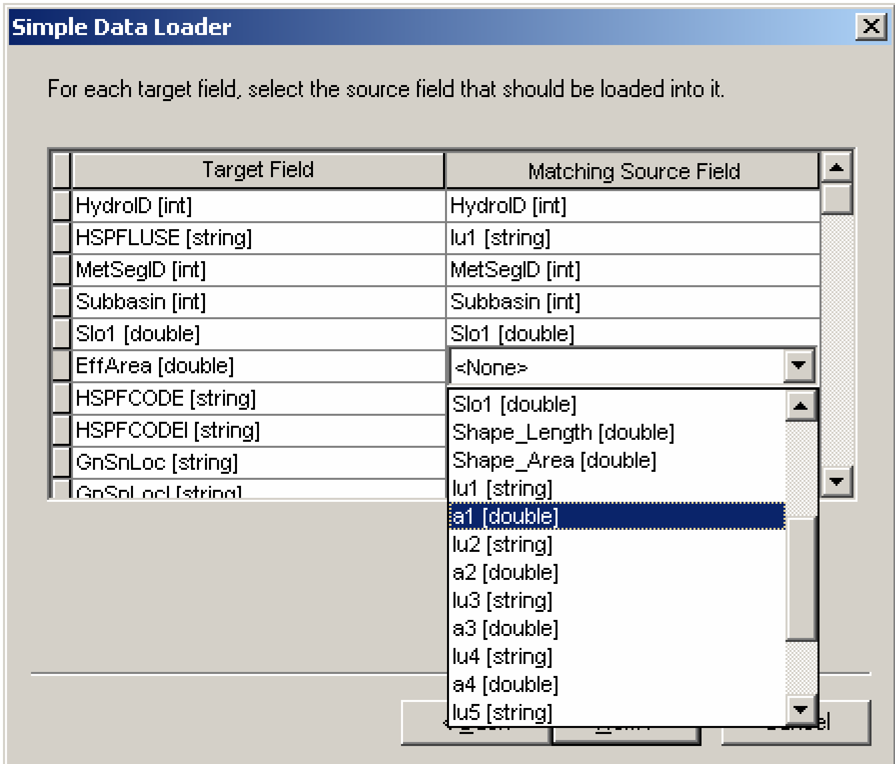


Figure A.4 Load data into the HSPFLandSegment feature class: Landuse Areas

(11) The HSPFLandSegment feature class should now contain features representing the amount of the first type of landuse contributing to each of the River Segments. Repeat steps 8-10 for the remaining landuse types making sure to load a set of Catchments for each type of landuse to be simulated, and to choose the correct fields each time.

Next Step:

With the HSPFLandSegment populated, the tools to extract information to intermediate text files can be used as with Polygon landuse data. The only difference is that the 'EffArea' attribute should now be used rather than the 'Shape_Area.' The actual shape associated with the HSPFLandSegment feature classes is no longer representative of the area of land represented by the feature.

Appendix B: NEXRAD Data to Mean Precipitation

Contents:

This appendix contains a brief description of the steps used to prepare NEXRAD data for use in the ArcGIS Timeseries Preprocessing methodology. *Note: This is not intended to be a detailed presentation of a robust methodology, but some prototype tools have been developed. The author of this thesis should be contacted if more information is desired.*

Objective:

The ultimate objective of the following tasks is to estimate the mean precipitation over polygons (GISMetSegment) and store the information in Arc Hydro timeseries format.

Tasks:

The two major tasks implemented to estimate mean precipitation over polyons are:

(1) Extract precipitation values from .grib files and write them to Arc Hydro format.
(2) Calculate the mean precipitation over polygons (GISMetSegment) using SQL queries on the data in the database

Specific Steps:

Once data is obtained from the NWS and uncompressed to a directory on the hard disk, code developed at the Texas Advanced Computing Center (TACC) and CRWR is used to:

extract from grib file

(1) Open a single grib file and read a selected set of values into computer memory
(2) Write values to an Arc Hydro timeseries table in a geodatabase attached to a feature class representation of the NEXRAD grid cells

(3) Clear program memory
(4) Repeat steps 1-3 for many time steps to obtain a series of values for each grid cell

calculate mean precipitation over polygons

(5) Intersect the polygon areas (GISMetSegment) with the NEXRAD grid cells polygon feature class
(6) Using a SQL query, calculate the mean precipitation over a single polygon (one MetSegment) for the entire length of record
(7) Write this value to Arc Hydro timeseries format now attached to the polygon feature (MetSegment) as mean precipitation
(8) Repeat steps 6-7 for each feature in the polygon feature class (GISMetSegment)

Next Step:

With mean precipitation attached to the features of the GISMetSegment feature class, tools from the ArcGIS Timeseries preprocessing system can be used to write .wdm files for modeling and update the .uci file to read these new .wdm datasets.

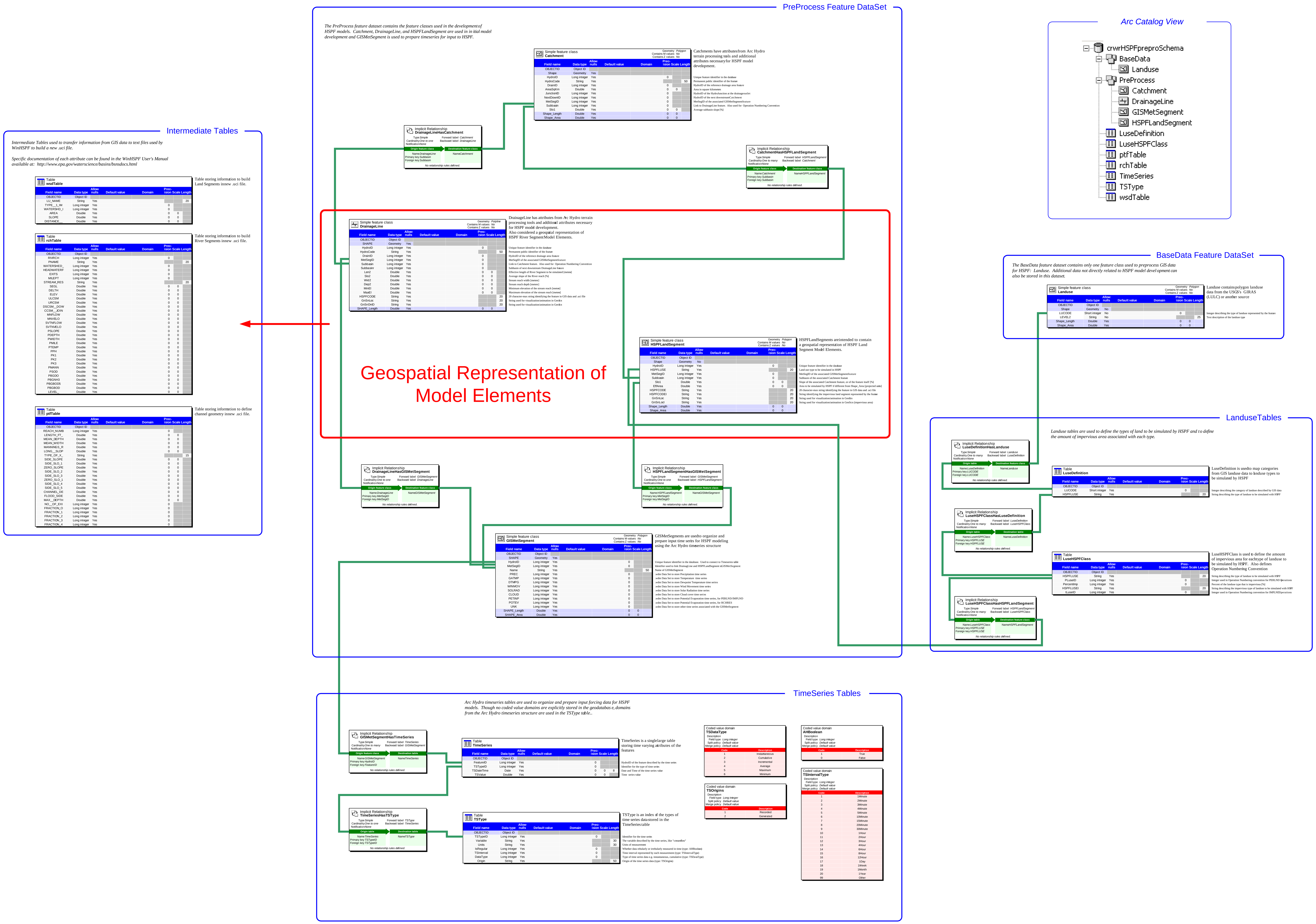
Appendix C – Diagram of Geodatabase Structure

ArcGIS HSPF and Timeseries Preprocessing Geodatabase

Center for Research in Water Resources

The University of Texas at Austin

Nate Johnson
Masters Thesis
2005



Appendix D: Attributes in Intermediate Text Files

This table presents the attributes of the intermediate text files and how they are calculated with the ArcGIS HSPF Preprocessing methodology. The color coding scheme is shown in Table D.1, and Table D.2 give each attribute from the three text files used in the methodology. The .psr file must be present to use the WinHSPF 'Create Project' tool, but does not contain any data extracted from GIS data.

Table D.1 Color coding scheme.

	default, value known, just written to the correct attribute
	value not used by WinHSPF for simple build .uci
	given, almost directly from the ArcHydro files (like length, names, etc…)
	given, and redundant, copy from another attribute
	needs to be calculated, and requires other data (like DEM, landuse to do it…)
	needs to be calculated, but redundant. copy from another attribute
	no easy way to get this value. Assign default if not available from another source
	no easy way to get this value, and redundant. Copy from another attribute

Table D.2 Attributes of intermediate text files.

Table	***Field Name***	***Source***	***Attribute***	***Method***
.wsd	LU Name	Landuse / LuseDefinition	HSPFLUSE	Join Landuse.LUCODE = LuseDefinition.LUCODE. Copy HSPFLUSE.
.wsd	Type (1=Imp, 2=Per)	LuseHSPFClass	PercImp	Join HSPFLandSegment.HSPFLUSE = LuseHSPFClass.HSPFLUSE. If PercImp > 0, create one Pervious Land Seg and one Pervious Land Seg
.wsd	Watershed-ID	Catchment	Subbasin	Copy Subbasin #

Table	Field Name	Source	Attribute	Method
.wsd	Area [acres]	HSPFLandSegment / LuseHSPFClass	Shape_Area / PercImp	Join HSPFLandSegment.HSPFLUSE = LuseHSPFClass.HSPFLUSE. If PercImp > 0, create one Pervious Land Seg, one Pervious Land Seg. Pervious Area = Shape_Area (1 - PercImp) Impervious Area = Shape_Area * PercImp
.wsd	Slope [%]	Subbasin	Slo1	Copy Slo1
.wsd	Distance [ft]	Subbasin	Slo1	WinHSPF assigns based on Slope. If < 0.005, = 500. if < 0.01, = 400. if < 0.03, = 350. if < 0.07, 300. if < 0.1, = 250. if < 0.15, 200, else, 150.
.rch	RivRch	DrainageLine	Subbasin	Copy Subbasin #
.rch	Pname	DrainageLine	Subbasin	Defaults to "Stream " & <Subbasin #>
.rch	Watershed-ID	DrainageLine	Subbain	Copy Subbasin #
.rch	HeadwaterFlag	DrainageLine	Subbasin / Subbasinr	Loop through DrainageLine features to see if any other DrainageLine has this feature as it's 'downstream' DrainageLine
.rch	Exits	N/A	N/A	Defaults to 1
.rch	Milept	N/A	N/A	N/A
.rch	Stream/Reservoir Type	N/A	N/A	Defaults to S
.rch	Segl [miles]	DrainageLine	Shape_Length / Len2	Copy Shape_Length or Len2. Convert from projected units
.rch	Delth [feet]	DrainageLine	MaxEl / MinEl	Calculate difference of MaxEl - MinEl
.rch	Elev [feet]	DrainageLine	MaxEl / MinEl	Calculate average of MaxEl and MinEl
.rch	Ulcsm	N/A	N/A	N/A
.rch	Urcsm	N/A	N/A	N/A

Table	***Field Name***	***Source***	***Attribute***	***Method***
.rch	Dscsm (downstream)	DrainageLine	Subbasinr	Copy Subbasinr #
.rch	Ccsm (joins with and down)	N/A	N/A	N/A
.rch	Mnflow	N/A	N/A	N/A
.rch	Mnvelo	N/A	N/A	N/A
.rch	Svtnflow	N/A	N/A	N/A
.rch	Svtnvelo	N/A	N/A	N/A
.rch	Pslope	DrainageLine	Slo2	Copy Slo2
.rch	Pdepth	DrainageLine	Wid2	Copy Wid2, if not > 0 then default to 30 ft
.rch	Pwidth	DrainageLine	Dep2	Copy Dep2, if not > 0 then default to 5 ft
.rch	Pmile	N/A	N/A	N/A
.rch	Ptemp	N/A	N/A	N/A
.rch	Pph	N/A	N/A	N/A
.rch	*Pk1*	N/A	N/A	N/A
.rch	*Pk2*	N/A	N/A	N/A
.rch	*Pk3*	N/A	N/A	N/A
.rch	*Pmann*	N/A	N/A	N/A
.rch	Psod	N/A	N/A	N/A
.rch	Pbgdo	N/A	N/A	N/A
.rch	Pbgnh3	N/A	N/A	N/A
.rch	Pbgbod5	N/A	N/A	N/A
.rch	Pbgbod	N/A	N/A	N/A
.rch	Level	N/A	N/A	N/A
.ptf	Reach Number	DrainageLine	Subbasin	Copy Subbasin #

Table	***Field Name***	***Source***	***Attribute***	***Method***
.ptf	Length(ft)	DrainageLine	Shpe_Length / Len2	Copy Shape_Length or Len2. Convert from projected units
.ptf	Mean Depth(ft)	DrainageLine	Dep2	Copy Dep2, if not > 0 then default to 5 ft
.ptf	Mean Width (ft)	DrainageLine	Wid2	Copy Wid2, if not > 0 then default to 30 ft
.ptf	Mannings Roughness Coeff.	N/A	N/A	Defaults to 0.0015
.ptf	Long. Slope	DrainageLine	Slo2	Copy Slo2
.ptf	Type of x-section	N/A	N/A	Defaults to 'Trapezoidal'
.ptf	Side slope of upper FP left	N/A	N/A	defaulted to 0.5
.ptf	Side slope of lower FP left	N/A	N/A	defaulted to 0.5
.ptf	Zero slope FP width left(ft)	DrainageLine	Wid2	Copy Wid2, if not > 0 then default to 9.14 ft
.ptf	Side slope of channel left	N/A	N/A	defaulted to 1.0
.ptf	Side slope of channel right	N/A	N/A	defaulted to 1.0
.ptf	Zero slope FP width right(ft)	DrainageLine	Wid2	Copy Dep2, if not > 0 then default to 1.52
.ptf	Side slope lower FP right	N/A	N/A	defaulted to 0.5
.ptf	Side slope upper FP right	N/A	N/A	defaulted to 0.5
.ptf	Channel Depth(ft)	DrainageLine	Dep2	Mean Depth times 1.25

Table	***Field Name***	***Source***	***Attribute***	***Method***
.ptf	Flood side slope change at depth	DrainageLine	Dep2	Mean Depth times 1.875
.ptf	Max. depth	DrainageLine	Dep2	Mean Depth times 62.5
.ptf	No. of exits	N/A	N/A	defaulted to 1
.ptf	Fraction of flow through exit 1	N/A	N/A	defaulted to 1.0
.ptf	Fraction of flow through exit 2	N/A	N/A	defaulted to 1.0
.ptf	Fraction of flow through exit 3	N/A	N/A	defaulted to 1.0
.ptf	Fraction of flow through exit 4	N/A	N/A	defaulted to 1.0
.ptf	Fraction of flow through exit 5	N/A	N/A	defaulted to 1.0

Bibliography

AQUA TERRA Consultants. AQUA TERRA Consultants: Specialists in Water Resources, Computer Modeling, and Environmental Assessment. 8 Aug. 2005 <www.hspf.com>.

Bhaskar, Nageshwar R., Wesley P. James, and Ravikumar S. Devulapalli. "Hydrologic Parameter Estimateion Using Geographic Information Sysgem." Journal of Water Resources Planning and Management. 118(5) (1992): 492-512.

Bicknell, Brian R., et. al. Hydrologic Simulation Program in Fortran. v.12 User's Manual. 2001. USEPA, Athens, GA.

CUAHSI. CUAHSI Background. 9 Aug. 2005 <http://www.cuahsi.org/about/history.html>.

DeVantier and Feldman. "Review of GIS Applications in Hydrologic Modeling." Jour. Water Resources Planning and Management ASCE. 119(2) (1993): 246-261.

Duda, Paul., J. Kittle, Jr., M. Gray, P. Hummel, R. Dusenbury. An Interactive Windows Interface to HSPF (WinHSPF). 2001 <http://www.epa.gov/waterscience/basins/bsnsdocs.html>.

Duda, Paul, et. al. "BASINS 4.0 –Flexible Integration of Components and Data for Watershed Assessment and TMDL Development." Unpublished. AQUA TERRA Consultants, Decatur, GA.

EMRL (Environmental Modeling Research Laboratory). WMS. 8 Aub. 2005 <http://www.emrl.byu.edu/wms.htm>.

EPA. "BASINS Technical Note 6." Office of Water. 2000. EPA-823-R00-012 <http://www.epa.gov/waterscience/basins/tecnote6.pdf>

EPA. "Better Assessment Science Integrating point and Nonpoing Sources. User's Manual." Office of Water. 2001. EPA-832-B-01-001.

EPA. Water Quality Models. 9 Aug. 2005a <http://www.epa.gov/waterscience/wqm/>.

EPA. BASINS: Better Assessment Science Integrating Point & Nonpoint Sources. 8 Aug. 2005b <http://www.epa.gov/OST/BASINS/>.

EPA. WCS: Watershed Characterization System. 8 Aug. 2005c <http://wcs.tetratech-ffx.com/>.

EPA. Land Use/Land Cover GIRAS Spatial Data in the Conterminous U.S. 9 Aug. 2005d <http://www.epa.gov/nsdi/projects/giras.htm>.

ESRI. Developing for ArcGIS. 9 Aug. 2005a <http://edndoc.esri.com/arcobjects/9.0/default.asp?URL=/arcobjects/9.0/ArcGISDevHelp/ArcGIS%20Product%20Line/Developing_StartPage.htm>.

ESRI. Geoprocessing Library Overview. 9 Aug. 2005b <http://edndoc.esri.com/arcobjects/9.0/default.asp?url=/arcobjects/9.0/ComponentHelp/esriGeoprocessing/Geoprocessing_overview.htm>.

ESRI. Desktop GIS. 9 Aug. 2005c <http://www.esri.com/software/arcgis/about/desktop.html>.

Garbrecht et. al. "GIS and distributed watershed models. 1: Data coverages and sources." Journal of Hydrologic Engineering. 6(6) (2001): 608-614.

Gopalan, Hema and David R. Maidment. "WRAPHydro Data Model: Finding Input Parameters for the Water Rights Analysis Package." Thesis. The University of Texas at Austin, 2003. CRWR Online Report 03-03. <http://www.crwr.utexas.edu/reports/2003/rpt03-3.shtml>.

Gueudet, Pierre, and David R. Maidment. HEC-PREPRO: "The Influence of Post-Spacing Density of DEMs Derived from LIDAR on Flood Modeling." Thesis.

The University of Texas at Austin, 2004. CRWR Online Report 04-7. <http://www.crwr.utexas.edu/reports/2004/rpt04-7.shtml>.

Hellweger, Ferdinand and David R. Maidment. "HEC-PREPRO: A GIS Preprocessor for Lumped Parameter Hydrologic Modeling Programs." Thesis. The University of Texas at Austin, 2003. CRWR Online Report 97-8. <http://www.crwr.utexas.edu/reports/2003/rpt03-3.shtml>.

HydroComp. Hydrologic Simulation Models: An Overview. 16 Jul. 2005 <http://www.hydrocomp.com/simoverview.html>.

Maidment, David R. Applied Hydrology. New York. McGraw Hill, 1988.

Maidment, D. R. "GIS and hydrologic modeling." Proceedomgs, 1st Int. Symposium / Workshop on GIS and Envir. Modeling. (1991).

Maidment, David R., ed. Arc Hydro: GIS for Water Resources. Redlands, CA. ESRI Press. 2002.

Neary, V.S., E. Habib, and M. Fleming. "Hydrologic Modeling with NEXRAD Precipitation in Middle Tennessee." Journal of Hydrologic Engineering. 9(4) (2004): 339-349.

NOAA. West Gulf River Forecast Center. National Weather Service. 8 Aug. 2005a <http://www.srh.noaa.gov/wgrfc/>.

NOAA. Archive of River Forecast Center Operational NEXRAD Data. NOAA Hydrologic Data Systems Group. 8 Aug. 2005b <http://dipper.nws.noaa.gov/hdsb/data/nexrad/nexrad.html>

NRCS. Soil Survey Geographic (SSURGO) Database. 9 Aug. 2005a <http://www.ncgc.nrcs.usda.gov/products/datasets/ssurgo/index.html>.

NRCS. State Soil Geographic (STATSGO) Database. 9 Aug. 2005b <http://www.ncgc.nrcs.usda.gov/products/datasets/statsgo/>.

Obenour, Dan R and David R. Maidment. "Arc Hydro Developments for the Lower Colorado River Basin." Thesis. The University of Texas at Austin, 2004. CRWR Online Report 04-06. <http://www.crwr.utexas.edu/reports/2004/rpt04-6.shtml>.

Ogden et. al. "GIS and distributed watershed models. II: Modules, interfaces and models." Journal of Hydrologic Engineering. 6(6) (2001): 515-523.

Olivera, Francisco and David R. Maidment. GIS Tools for HMS Modeling Support. 19th Annual ESRI Users Conference, 1999. 15 Jul. 2005 <http://gis.esri.com/library/userconf/proc99/proceed/papers/pap801/p801.htm>.

Olivera, Francisco. "Personal Contact." 8 Aug. 2005. Website: <http://ceprofs.tamu.edu/folivera/>

Roboyo, Oscar and David R. Maidment. Map to Map: Converting a NEXRAD Rainfall Map into a Flood Inundation Map. Dissertation. The University of Texas at Austin, 2005. CRWR Online Report 05-01. <http://www.crwr.utexas.edu/reports/2005/rpt05-1.shtml>.

SARA (San Antonio River Authority). Regional Watershed Management. 8 Aug. 2005. <http://www.sara-tx.org/site/flood_control/reg_watersheed_mngt.html>.

Singh, Vijay P., and David A. Woolhiser. "Mathematical Modeling of Watershed Hydrology." Journal of Hydrologic Engineering. 7 (4): (2002): 270-292.

Unidata. "The NetCDF Users' Guide." Unidata Programming Center. January 2005.

USACE. Geospatial Modeling Extension: HEC-GeoHMS. 8 Aug. 2005a <http://www.hec.usace.army.mil/software/hec-hms/hechms-geohms.html>.

USACE. HEC-GeoRAS. 8 Aug. 2005b <http://www.hec.usace.army.mil/software/hec-ras/hecras-hec_georas.html>.

USACE. HEC Data Storage System. Visual Utility Engine User's Manual. May 2005c. Report CPD-79. <http://www.hec.usace.army.mil/software/hec-dss/hecdss-dss.html>.

USGS. User's Manual for ANNIE, Version 2, a Computer Program for InteractiveHydrologic Data Management. Reston, VA. 1995. Report 95-4085.

USGS. A Tool for the GENeration and Analysis of Model Simulateion Scenarios for Watersheds (GenScn). Water Resources Investigations. 1998. Report 98-4134.

USGS. National Elevation Dataset. 9 Aug. 2005 <http://ned.usgs.gov/>.

USGS and EPA. National Hydrography Dataset. 9 Aug. 2005 <http://nhd.usgs.gov/>.

USGS, EPA, and NOAA. Multi-Resolution Land Characteristics Consortium. 9 Aug. 2005 <http://www.mrlc.gov/index.asp>.

Whiteaker, Tim and David R. Maidment. Geographically Integrated Hydrologic Modeling Systems. Dissertation. The University of Texas at Austin, 2004. CRWR Online Report 04-04. <http://www.crwr.utexas.edu/reports/2004/rpt04-4.shtml>.

WMO. A Guide to the Code Form FM 92-IX Ext. GRIB. 15 Jul. 2005 <http://www.wmo.ch/web/www/WDM/Guides/Guide-binary-2.html>.

Wurbs. Water Rights Analysis Package (WRAP) Modeling System Users Manual. Texas Water Resources Institute. 2003. TR-256.

Xie, Hongjie, et. al. "GIS-based NEXRAD Stage III precipitation database: automated approaches for data processing and visualization." Computers and Geosciences. (2005): 65-76.

Zhang, Z. Y. et al. "Use of next generation weather radar data and basin disaggregation to improve continuous hydrograph simulations." Journal of Hydrologic Engineering. 9(2) (2004): 103-115.

Vita

Nathan William Johnson was born in Brainerd, MN on September 3, 1980, the son of Owen and Nancy Johnson. After completing his work at Pine River/Backus High School, he enrolled at Princeton University in Princeton, NJ. He received his Bachellor of Science degree in Civil and Environmental Engineering from Princeton University in June, 2003. In September, 2003 he began graduate studies at The University of Texas at Austin.